Fall 12-30-2020

High Fidelity Computational Modeling and Analysis of Voice Production

Weili Jiang
weili.jiang@maine.edu

HIGH FIDELITY COMPUTATIONAL MODELING AND ANALYSIS OF VOICE PRODUCTION

By

Weili Jiang

B.S. Southeast University, 2012

M.S. Southeast University, 2015

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Mechanical Engineering)

The Graduate School

The University of Maine

Aug 2020

Advisory Committee:

Xudong Zheng, Associate Professor of Mechanical Engineering, Advisor

Qian Xue, Assistant Professor of Mechanical Engineering, Co-Advisor

Senthil S. Vel, Arthur O. Willey Professor of Mechanical Engineering of Mechanical Engineering

Kimberly Huguenard, Assistant Professor of Civil and Environmental Engineering

Zhaoyan Zhang, Professor of Department of Head and Neck Surgery, David Geffen School of Medicine at UCLA

HIGH FIDELITY COMPUTATIONAL MODELING AND ANALYSIS OF VOICE PRODUCTION

By Weili Jiang

Dissertation Advisor: Dr. Xudong Zheng

An Abstract of the Dissertation Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Mechanical Engineering)
December 2020

This research aims to improve the fundamental understanding of the multiphysics nature of voice production, particularly, the dynamic couplings among glottal flow, vocal fold vibration and airway acoustics through high-fidelity computational modeling and simulations. Built upon in-house numerical solvers, including an immersed-boundary-method based incompressible flow solver, a finite element method based solid mechanics solver and a hydrodynamic/aerodynamic splitting method based acoustics solver, a fully coupled, continuum mechanics based fluid-structure-acoustics interaction model was developed to simulate the flow-induced vocal fold vibrations and sound production in birds and mammals. Extensive validations of the model were conducted by comparing to excised syringeal and laryngeal experiments. The results showed that, driven by realistic representations of physiology and experimental conditions, including the geometries, material properties and boundary conditions, the model had an excellent agreement with the experiments on the vocal fold vibration patterns, acoustics and intraglottal flow dynamics, demonstrating that the model is able to reproduce realistic phonatory dynamics during voice production. The model was then utilized to investigate the effect of vocal fold inner structures on voice production. Assuming the human vocal fold to be a three-layer structure, this research focused on the effect of longitudinal variation of layer thickness as well as the cover-body thickness ratio on vocal fold vibrations. The results showed that the longitudinal variation of the cover

and ligament layers thicknesses had little effect on the flow rate, vocal fold vibration amplitude and pattern but affected the glottal angle in different coronal planes, which also influenced the energy transfer between glottal flow and the vocal fold. The cover-body thickness ratio had a complex nonlinear effect on the vocal fold vibration and voice production. Increasing the cover-body thickness ratio promoted the excitation of the wave-type modes of the vocal fold, which were also higher-eigenfrequency modes, driving the vibrations to higher frequencies. This has created complex nonlinear bifurcations. The results from the research has important clinical implications on voice disorder diagnosis and treatment as voice disorders are often associated with mechanical status changes of the vocal fold tissues and their treatment often focus on restoring the mechanical status of the vocal folds.

DEDICATION

To my family and my best friend …

ACKNOWLEDGEMENTS

There are a number of people I would like to thank for offering their support during my PhD study. First and foremost, I would like to thank my advisor and co-advisor Prof. Xudong Zheng and Prof. Qian Xue in Mechanical Engineering department in University of Maine. Without their guidance, encourage and help, I cannot achieve the current progress. I also want to thank Dr. Sid M. Khosla, Charles Farbos de Luzan, Liran Oren from University of Cincinnati and Prof. Coen P.H. Elemans and Jeppe H. Rasmussen from University of Southern Denmark for collaboration and providing important experiment support.

I would also like to thank other colleagues in our group, Xiaojian Wang, Dariush Dodaghi, Mohammadreza Movahhedi, Geng Liu, Yang Zhang, for their help and suggestion and my committee members for reading my dissertation and providing me with important suggestions.

I will also thank my family members: my parents and Biao Geng. Their love always encourages me.

Finally, I would like to thank the financial support from National Institute on Deafness and Other Communication Disorders (NIDCD) through Grant No.1R03DC014562, No. R01DC009435 and the Janet Waldron Doctoral Research Fellowship from the University of Maine, and the computational resource from Extreme Science and Engineering Discovery Environment (XSEDE) through allocation Award Nos. TG-BIO150055 and TG-CTS180004) and the advanced computing group from the University of Maine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

E_p	Transverse Young's modulus
E_{pz}	Longitudinal Young's modulus
v_p	Transverse Poisson ratio
v_{pz}	Longitudinal Poisson ratio
G_{pz}	Longitudinal shear modulus
F_0	Fundamental frequency
Q_{mean}	Mean flow rate
Q_{max}	Maximum flow rate
τ_o	Open quotient
τ_s	Skewness quotient
P_{SG}	Subglottal pressure
ρ, ρ_o, ρ'	Total, incompressible, compressible perturbed density
U, u, u'	Total, incompressible, compressible perturbed velocity
P, p, p'	Total, incompressible, compressible perturbed pressure

CHAPTER 1 INTRODUCTION

1.1. Physiology and anatomy

The sound production system in the human body generally consists of the lung, the larynx, the vocal tract, the pharynx, the oral and nasal cavities (Titze, 2000). Briefly, neural commands are sent from the brain to contract the laryngeal muscles. Through muscle contraction, the geometry and tension of the laryngeal tissues could be adjusted to specific phonation condition. This process is known as posturing. Lung pressure, which is around 1 kPa in normal condition, functions to press the air to flow through the larynx. When pressed together during posturing (adducted), such as in the production of voiced sounds like the vowels, the vocal folds would vibrate due to lung pressure and generate pulsatile glottal flow in the frequency around 100-200Hz in adults. When not fully adducted during posturing, such as the production of unvoiced sound like /h/, the vocal folds would vibrate in very small amplitude and are never fully closed. The oscillating glottal flow would be filtered by the supraglottal tract, in which the energy in certain frequencies would be amplified mainly according to the shape of the tract. By changing the position of the tongue, teeth, lips, and the shape of the oral cavity etc., the shape of the tract is altered, resulting in different formant frequencies. Different combinations of the formant frequencies form the perceptive bases of different vowels and consonants.

The larynx is a key component in this process, responsible for the generation of the pulsatile glottal flow, which is the sound source. It is located just above the trachea, consisting mainly of cartilages, muscles, and the vocal folds. The larynx is in the order of 1cm in size and is highly flexible without rigid bony connection to the skeleton (Titze, 2000). Other than phonation, the more vital role of the larynx is protecting the lower airway from foreign bodies such as food. During swallowing, the

epiglottis flaps down to block the entrance of the larynx and the vocal fold would also tightly close, preventing food dropping into the larynx and lung.

1.1.1 Cartilages and muscles

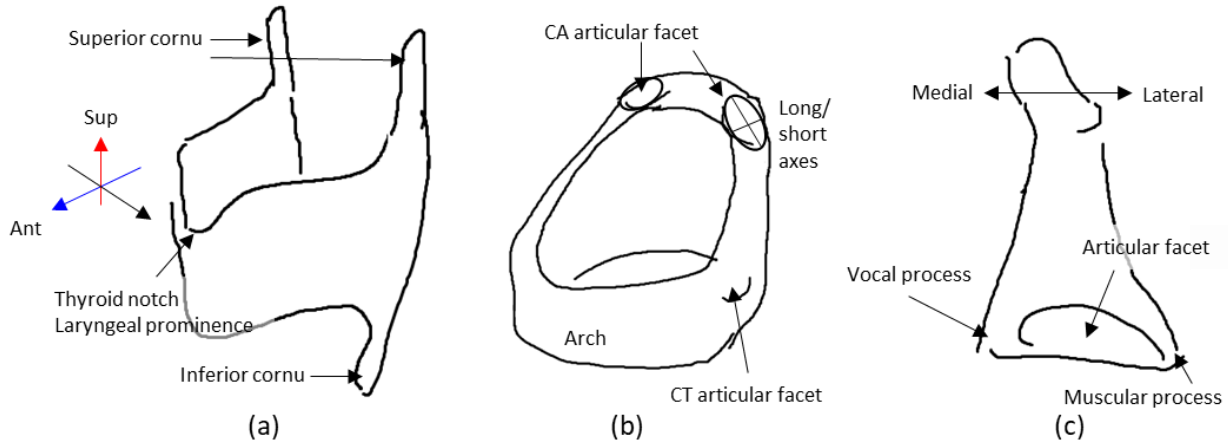


Figure 1.1 The laryngeal cartilages. (a) The thyroid, (b) the cricoid and (c) the arytenoid cartilages.

The laryngeal cartilages include the cricoid cartilage, the thyroid cartilage and one pair of the arytenoid cartilages (Figure 1.1). The thyroid cartilage (Figure 1.1(a)) consists of two plates that joined at the anterior side with an angle of 90-120 degrees. There are four projections, known as the superior cornua and inferior cornua. The inferior cornua joins the cricoid cartilage at the cricothyroid (CT) articular facet and the superior cornua connects to the hyoid bone. The cricoid cartilage (Figure 1.1(b)) is located below the thyroid cartilage, atop the tracheal rings. It can rotate forwardly and anteriorly toward the thyroid cartilage around the CT joint (CT articular facet). The cricoarytenoid (CA) articular facets are small smooth elliptic surfaces that the arytenoid cartilage would attach to. The arytenoid cartilages (Figure 1.1(c)) are roughly pyramidal in shape. It is attached to the cricoid cartilage through the CA articular facet, allowing rotation and translation motion. It can move in both medial-lateral and anterior-posterior direction. The vocal process of the arytenoid is the attachment of the vocal ligament.

The muscle processes are the attachment of the posterior cricoarytenoid (PCA) and lateral cricoarytenoid (LCA) muscles. The cricoid and thyroid cartilages contact on CT joint, which allows rotation and gliding (Figure 1.2(a)). The gliding can contribute 30-40% of the vocal fold elongation (Alipour-Haghighi, Perlman and Titze, 1991). The cricoid and arytenoid cartilages contact on CA joint, which allows the arytenoid cartilage to slide along and rotate about the long axis of the cricoid articular facet (Zhang, 2016b).

The intrinsic muscles in the larynx are the CT muscle, the thyroarytenoid (TA) muscle, the LCA muscle, the PCA muscle and the interarytenoid (IA) muscle. Different from the extrinsic muscles that move the location of the whole larynx, the intrinsic muscles are responsible for voice control through the adjustment of vocal fold shape and tension. CT-TA and PCA-LCA/IA are the two agonist-antagonist muscle pairs, responsible for adjusting the length and opening of the vocal folds, respectively.

The CT and TA muscles are the primary pitch control muscles (Figure 1.2 (a, b)). The CT muscle, including the rectus and oblique parts, connects the cricoid and the thyroid cartilages. Originating from the anterior arch of the cricoid, the CT rectus inserts upward to the lower border of the thyroid cartilage and the CT oblique inserts upward and backward to the inferior cornua of the thyroid cartilage. By contracting CT muscle, the cricoid arch could be elevated and the thyroid depressed due to the rectus part (Figure 1.2(b)). The CT oblique might contribute to the forward motion of the thyroid cartilage but its exact function is controversial (Ludlow, 2005). CT contraction increases vocal fold length and tension, resulting in higher pitch. The TA muscle connects the thyroid cartilage (below thyroid notch) and the arytenoid cartilage. Contrary to the CT muscle, TA contraction draws the arytenoid cartilages forward, which would shorten and thicken the vocal fold. It will stiffen the inner layers and slacken the cover layer. The effect of the TA contraction on the pitch depends on whether and to what extent is the TA

muscle involved in the vibration. If the vibration is deep into the TA muscle, the pitch might increase. If the vibration is mainly in the slackened cover layer, the pitch might decrease (Titze, 1988; Titze, Jiang and Drucker, 1988; Chhetri and Park, 2016).

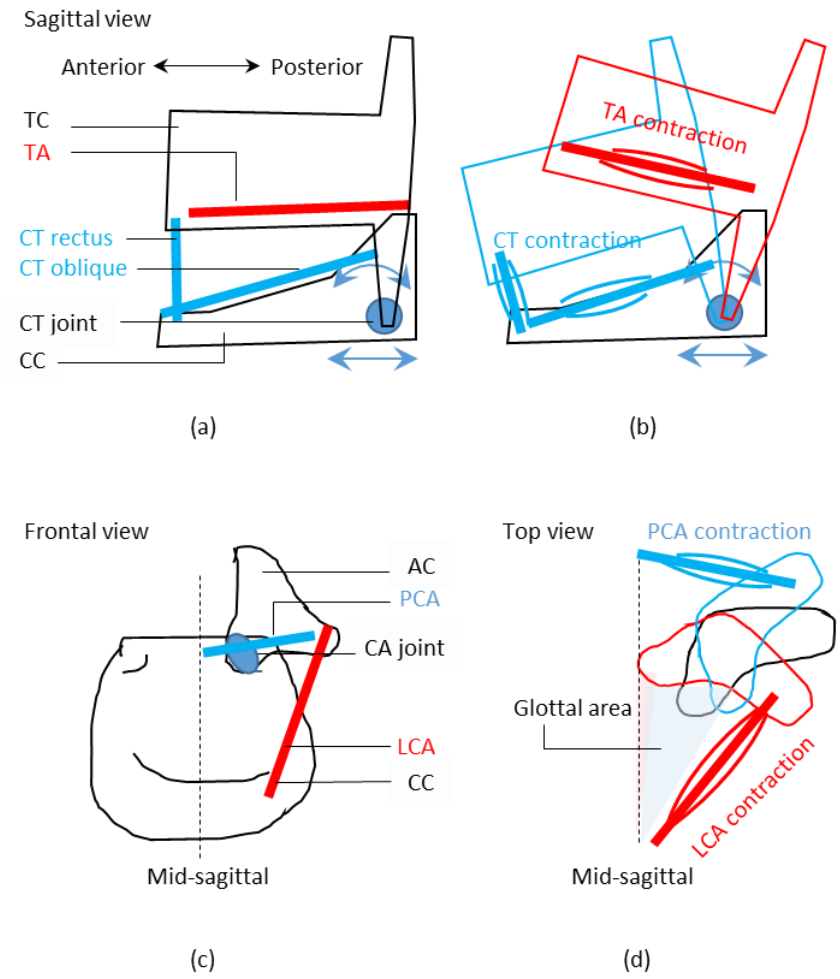


Figure 1.2 Main cartilages and muscles in the larynx. CC: cricoid cartilage; TC: thyroid cartilage; AC: arytenoid cartilage; CT: cricothyroid muscle; TA: thyroarytenoid muscle; LCA: lateral cricoarytenoid muscle; PCA: posterior cricoarytenoid muscle.

PCA is the primary abductor which moves the vocal folds laterally while the LCA/IA are the adductors that draw them medially (Figure 1.2 (c,d)). PCA originates from the posterior surface of the

cricoid cartilage and inserts upward and laterally to the muscular process of arytenoid cartilage. LCA originates from the cricoid cartilage and inserts into the muscular process of the arytenoid. IA connects the two arytenoid cartilages. The contraction of the PCA rotates and glides the arytenoid cartilage to move the vocal folds laterally at the posterior end. The contraction of the LCA draws the arytenoid forward and medially. It brings together vocal process thus also moves vocal folds medially at the posterior end. Contraction of the IA adducts the vocal folds by moving the two arytenoids medially, closing more in the posterior part of the glottis than in the anterior part. The function of abductors allows more air flowing through the glottis, which is important in unvoiced sound and breathing. The function of adductors closes the vocal folds which is critical in producing voiced sound.

Briefly, the intrinsic muscles are innervated by two main branches of the vagus nerve (Titze, 2000). One is the superior laryngeal branch that innervates the CT muscle only. The other branch is the recurrent laryngeal branch which innervates all the intrinsic muscles except the CT muscle.

1.1.2 Morphology of the soft tissue of the vocal folds

The vocal fold is a layered structure consisting of the epithelium, the lamina propria and the TA muscle (Hirano and Kakita, 1985; Titze, 2000). As shown in Figure 1.3, the epithelium is a very thin layer (0.05-0.08mm) made up of stratified squamous cells, with a Young's modulus around 100kPa which is much higher than the other layers (Hirano and Kakita, 1985). The lamina propria is non-muscular tissue that is usually divided into three layers according to the density of fibrous components (Hirano and Kakita, 1985). The superficial layer of the lamina propria consists of mainly loosely organized elastin fibers surrounded by interstitial fluids. It is about 0.5mm thick in the middle. In the intermediate layer, the elastin fibers are uniformly oriented in the anterior-posterior direction. The deep layer of the lamina propria mainly consists of collagen fibers aligned in the anterior-posterior direction. The intermediate

and deep layers of the lamina propria together are about 1-2mm thick (Titze, 2000). Below the deep layer of the lamina propria is the TA muscle which is about 7-8mm thick.

A lot of works have been done in measuring the material properties of the vocal fold layers (review papers: Miri, 2014; Dion *et al.*, 2016). Typical measurements including traction testing, rheometry and indentation are performed to obtain the stress-strain (force-elongation) relationship, shear modulus and local force-deformation measurement in small samples, respectively. Elastic properties of the vocal folds also vary locally. The elastic modulus of the middle part of the ligament can be much stiffer (~10 times) than the two ends (Kelleher *et al.*, 2010; Bhattacharya, Kelleher and Siegmund, 2015). It has also been reported that the inferior medial surface is stiffer than the superior medial surface(Chhetri, Zhang and Neubauer, 2011; Oren *et al.*, 2014).

In phonation studies, for simplicity, these five histological layers are usually grouped into two or three layers based on their elastic properties (Hirano and Kakita, 1985; Titze, 2000) (Figure 1.3). In the two-layer model, the cover layer consists of the epithelium, the superficial and intermediate layers of the lamina propria, and the body layer consists of the deep layer of the lamina propria and the TA muscle. In the three-layer model, the cover layer consists of the epithelium and the superficial layer of the lamina propria, the ligament layer consists of the intermediate and deep layers of the lamina propria, and the body layer consists of the TA muscle. Vocal fold vibration is the direct result of the interaction between glottal flow and the layered vocal fold. Both the two-layer model (Hirano, 1974; Titze, 1988; Cook, Nauman and Mongeau, 2008; Chhetri, Zhang and Neubauer, 2011; Zhang and Luu, 2012; Daily and Thomson, 2013) and the three-layer model (Titze and Talkin, 1979; Rosa *et al.*, 2003; Xue *et al.*, 2011; Zheng *et al.*, 2011; Weiss *et al.*, 2016) are widely used to study the effect of the structure and elasticity of vocal fold layers on the vibrations.

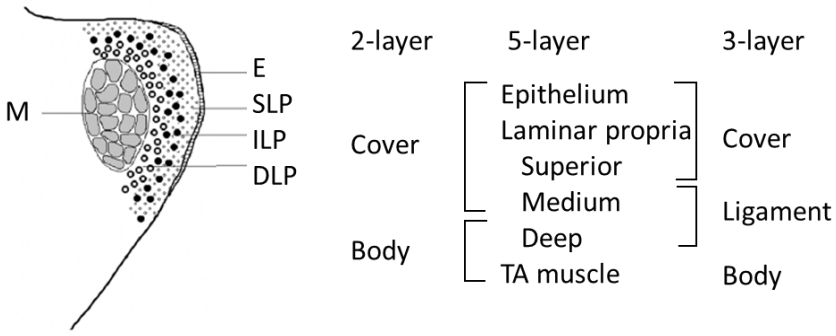


Figure 1.3 Layered structure of the vocal fold. E: Epithelium; SLP: superior layer of the lamina propria; ILP: intermediate layer of the lamina propria; DLP: deep layer of the lamina propria; M: TA muscle.

1.2 Physics of the sustained energy transfer

During vocal fold vibration, there is a continuous energy transfer from glottal air flow to the vocal fold tissue, avoiding the need of nerve stimulation at the same frequency of phonation frequency (Van den Berg, 1958). Asymmetric pressure during glottal opening and closing is responsible for the continuous energy transfer. One typical reason of such asymmetric pressure is the mucosal wave propagation (Titze, 1988). Figure 1.4 shows a typical pattern of the vocal fold vibration. The inferior side of the vocal fold always leads the motion, forming a convergent glottal shape in the opening phase and a divergent shape in the closing phase. The phase difference between the inferior and superior side of the vocal fold varies from 30^{o}/mm to 60^{o}/mm (Titze, Jiang and Hsiao, 1993). When the glottis is a convergent shape, intraglottal pressure is a positive value, decreasing from the subglottal pressure to ambient pressure in the streamwise direction. When the glottis is a divergent shape, due to flow separation, the intraglottal pressure is around 0 or a negative value in the glottis. Therefore, the direction of the pressure force applied on the vocal fold wall is the same with the vocal fold motion during most of the phase, resulting in continuous positive energy transfer from the air flow to tissue.

Such convergent-divergent type of vibration pattern is also observed experimentally in other mammals and birds (Elemans *et al.*, 2015). If the vocal folds open and close in the same shape, e.g. in the one-mass model, external mechanism such as an inertive supraglottal tract is necessary to generate the asymmetric pressure (Flanagan and Landgraf, 1968). The difference of aerodynamics of the glottal flow during flow increasing and decreasing, e.g. flow separation in the closing phase, could also generate asymmetric pressure distribution (Zhang, 2016b).

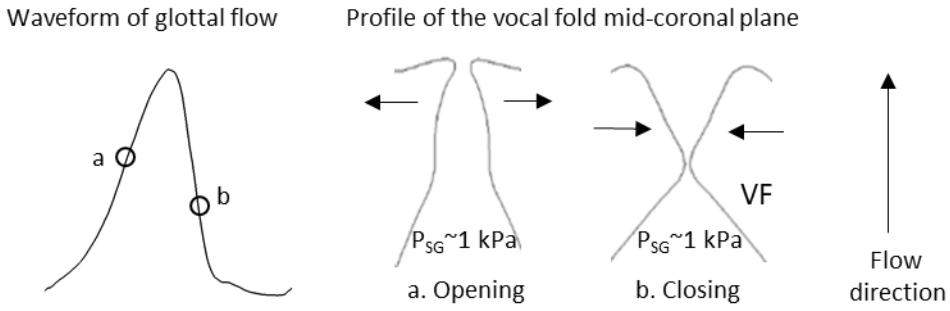


Figure 1.4 Mucosal wave propagation.

1.3 Modelling techniques

With the development of computer technology, numerical models of the voice production improved remarkably in the recent decades. Nowadays, both reduced order models and the partial differential equation based models are extensively used. Numerical modelling is an alternative approach to the experiment measurement when it is limited by the small size of the larynx and the accessibility in vivo (Gómez *et al.*, 2017).

Numerical modelling has some advantages over the experiment measurement. First, it can provide comprehensive three-dimensional quantifications of the vocal fold geometry and tension distribution. Second, it can calculate the active stress in the muscles while in most of the conditions of

the experiment measurement, the preparation of an excised larynx already removes the possibility of muscle stimulation (Garcia and Herbst, 2018). Third, parameters are better controlled. In the experiment, the adjustment of one parameter is usually accompanied with change in other parameters, making the analysis of the causal relationship even more complicated.

1.3.1 The tissue model

1.3.1.1 Reduced order model

In lumped-mass models, the vocal folds are simplified as mass-spring-damper systems that laterally connect to rigid walls. In the one-mass model (Flanagan and Landgraf, 1968), inertive vocal tract loading is necessary to achieve the self-sustained oscillation due to the lack of mucosal wave. In the two-mass model (Ishizaka and Flanagan, 1972a), the vocal fold is separated into a superior mass and an inferior mass. The two masses are connected through another spring, enabling the ability to capture the vertical phase difference. Hence the self-sustained oscillation could be obtained with or without vocal tract. With asymmetric parameters, two-mass model could also capture nonlinear features such as asymmetric and chaotic vibration (Steinecke and Herzel, 1995; Xue *et al.*, 2010). Variations based on the two-mass model include more mass units distributed in the sagittal plane (Titze, 1973; Wong *et al.*, 1991; Alipour *et al.*, 2011). However, the above models do not include the layered structure feature of the vocal fold. To improve this, (Story and Titze, 1995) developed the three-mass model based on the body-cover concept (Hirano, 1974). It allows more physiologically relevant laryngeal adjustments. For example, the contraction of the TA muscle could be reflected by adjusting the spring stiffness standing for the body layer, and the contraction of the CT muscle could be reflected by adjusting the spring stiffness in both cover and body layers. Another model based on the body-cover concept is the rotating plate model that the cover layer is represented by a rotating surface that can propagate mucosal wave

(Titze, 1988; Fulcher and Scherer, 2019). The mucosal wave speed is employed as the model input which is more experimentally measurable compared with the parameter of spring stiffness (Titze, 1988).

Reduced order model is computationally efficient and has been employed in improving the understanding of the basis of the self-oscillation process, laryngeal pathology and speech synthesis (Erath *et al.*, 2013). To determine the model input of the reduced order model, rules have been provided to convert muscle activities, including the CT, TA, LCA, PCA muscles, into physical quantities such as the changes in the vocal fold geometry and stiffness (Titze and Story, 2002). Neurophysiological fluctuations in the muscle activation could also be incorporated into this model to make the simulation more realistic (Manriquez *et al.*, 2019). Optimization method to determine the two-mass model parameters has been developed by minimizing the trajectory difference of left/right vocal folds between high-speed recordings and the simulation results (Döllinger *et al.*, 2002; Gómez *et al.*, 2017). However, the relationship between model input and the tissue property is still complicated. Moreover, the reduced order models usually ignore the vertical motion of the vocal fold (Titze and Story, 2002) and are not very helpful in the study of vocal fold posturing.

1.3.1.2 Continuum mechanics-based model

The continuum mechanics-based models were proposed as early as Titze and Talkin (1979) and Alipour-Haghighi and Titze (1985). They are now widely used due to the development of the computer technology (Alipour, Berry and Titze, 2000; Suh and Frankel, 2007; Zhang, Neubauer and Berry, 2007; Luo *et al.*, 2008; Zheng *et al.*, 2009; Šidlof, Zörner and Hüppe, 2013, 2014). The simulation of vocal fold vibration around specific posturing position usually takes the assumption of small-deformation and employs linear material properties (Hirano and Kakita, 1985; Titze, 1988; Cook and Mongeau, 2007; Tao and Jiang, 2007; Xue *et al.*, 2014; Zhang, 2015; Erath, Zañartu and Peterson, 2017). Such simplification is

based on the observation that the stress-strain relationship of the laryngeal muscles is roughly linear during vibration (Titze, 2006a). Contrarily, vocal fold posturing includes large deformation that the linear material assumption is not suitable. In this condition, nonlinear materials are employed. As is widely applied in material models of biological tissues, the tissue of vocal fold is usually assumed to be incompressible or nearly incompressible.

In the studies of vocal fold vibration, considering the histological feature of the tissue and the observation that vocal folds vibrate mainly in the transverse plane, the tissue, especially the muscle, is usually assumed to be transversely isotropic with a very high value of stiffness in the longitudinal direction (Berry, 2001; Cook, Nauman and Mongeau, 2008). The cover layer is assumed to be either isotropic (Zhang, 2014) or transversely isotropic (Xue *et al.*, 2014; Zhang, 2017a). The incompressible transversely isotropic material has only three independent parameters (E_p, E_{pz}, G_{pz}) and the isotropic material has only one parameter (E). These values can be measured experimentally through traction testing, rheometry and indention as mentioned in Section 1.1.2. Furthermore, as reported by (Cook, Nauman and Mongeau, 2008), based on a two-layer model, when the longitudinal to transverse stiffness ratio is greater than 30, the relative frequency difference is less than 5%. It indicates that E_{pz} could be assumed to be certain times (e.g. 30 times) larger than E_p to further reduce the material unknowns.

In the studies of vocal fold posturing, nonlinear material is required due to the large deformation. The muscle model, which is of especial importance in posturing process, is usually treated as a parallel connection of the passive and active components (Yin and Zhang, 2013; Smith and Hunter, 2014; Pham, Xue and Zheng, 2018). The passive stress could be calculated using isotropic (Hunter, Titze and Alipour, 2004), hyperelastic (Yin and Zhang, 2014, 2016), viscoelastic (Zhang, Siegmund and Chan, 2009; Smith and Hunter, 2014) or fiber-reinforced models (Holzapfel, Gasser and Ogden, 2000; Pham,

Xue and Zheng, 2018). The active component of the muscle stress depends on many factors such as the maximum muscle stress, fiber stretch, strain rate and the response time of the muscle. In Yin and Zhang (2013), the active stress was assumed to be related with the activation level and the optimal stretch, neglecting the time-dependence features. In Smith and Hunter (2014), the time-dependent features such as the strain-rate and the response time were also taken into consideration, with parameters determined by comparing with data derived from muscle stimulation experiments (Alipour-Haghighi, Perlman and Titze, 1991).

The advantages of the continuum mechanics-based model over the reduced order model are evident. It provides a direct relationship between tissue material and model parameters. Moreover, it has the ability to provide three-dimensional description of the tension distribution and the geometry even for the inner layers based on CT or MRI scans (Selbie, Gewalt and Ludlow, 2002; Chen *et al.*, 2012; Bakhshaee *et al.*, 2013). Continuum mechanics-based models have the advantage in studies with specific subjects such as clinic applications (Xue *et al.*, 2014; Chang *et al.*, 2016; Wu and Zhang, 2016), pathological conditions such as vocal fold nodules (Jiang, Diaz and Hanson, 1998; Gunter, 2004) and vocal fold posturing (Yin and Zhang, 2013, 2016).

However, compared with the reduced order models, the continuum mechanics-based models require more computational resource and time. To reduce the cost, based on the finding that the superposition of first several modes could describe the vocal fold vibration with high accuracy (Berry *et al.*, 1994), Zhang (2017b) proposed an eigenmode-based formulation of governing equation to greatly decrease the degrees of freedom of the governing equation to one hundred, while still having reasonable accuracy in parameters such as the fundamental frequency and the sound pressure level. This provides the potential to develop real-time applications such as speech synthesis.

1.3.2 The fluid model

1.3.2.1 One dimensional flow

The glottal flow could be modeled as simplified one dimensional, incompressible potential flow from the subglottal tract to the flow separation point while coupled with lumped-mass (Ishizaka and Flanagan, 1972a; Steinecke and Herzel, 1995) or continuum vocal fold models (Zhang, Neubauer and Berry, 2007). Corrections concerning the vena contracta effect at the glottal inlet, the viscous flow effect, the air inertance and the pressure recovery after the glottal exit (Ishizaka and Flanagan, 1972a) could be included in the model.

The location of flow separation is critical in the one-dimensional flow calculation when the vocal folds form a divergent shape. In some studies, the flow separation location is assumed to be fixed at the superior edge of the vocal fold (Zhang, Neubauer and Berry, 2007; Bailly *et al.*, 2008). Others take the location of the minimum cross-sectional area in the glottis (Steinecke and Herzel, 1995; Story and Titze, 1995) or slightly downstream of it where the area ratio to the minimum area is 1.2 -2 (Alipour and Scherer, 2004; Decker and Thomson, 2007; Zhang, 2009, 2014; Farahani and Zhang, 2014). Pelorson *et al.* (1994) and Cisonni *et al.* (2010) applied an approach that allows a moving separation point based on the boundary-layer theory. The different assumptions of the flow separation location affect the calculation of vocal fold loading which will further influence the vibration pattern (Farahani and Zhang, 2014).

The acoustic propagation in the supraglottal tract is expensive to model using direct numerical simulation. One popular approach employed in the one-dimensional models is the wave-reflection algorithm (Ishizaka and Flanagan, 1972a; Story and Titze, 1995; Titze, 2000). The supraglottal tract is discretized into serial-connected short cylindrical tubes with constant cross-section-area in each section.

Planar wave propagates in the tract. At the interface of the tubes, the reflection coefficient defined as the ratio of reflective pressure to the incident pressure can be calculated as a function of the cross-section-area ratio of the two sections with the assumption that the air density and the speed of sound are constants in the vocal tract (Titze, 2000). The boundary condition of the mouth is simplified as an open-end with the reflection coefficient to be -1. Therefore, the sound pressure transmitted from the vocal fold to the mouth can be calculated. The solution of the acoustic propagation could be coupled with the vocal fold vibration (Titze and Talkin, 1979; Zhang, 2015) or not (Steinecke and Herzel, 1995; Tao *et al.*, 2007; Elemans, Zaccarelli and Herzel, 2008) depending on the specific problem being studied.

With low computational cost, the one-dimensional flow model helps in understanding the physics of the self-sustained vibration, provides sound source in speech synthesis, and makes the parametric studies practical. Therefore, efforts have been made to increase the fidelity of the one-dimensional flow model including considering the flow asymmetry (Erath *et al.*, 2019), applying a momentum based flow model (Li *et al.*, 2020) and adding the viscous loss term predicted by deep neural network model (Zhang, Zheng and Xue, 2020).

1.3.2.2 Partial differential equation based flow

Partial differential equation (PDE) based flow models employ the Navier-Stokes equation to describe the flow field. They have the capability to provide flow features in two or three dimensions with stationary (Suh and Frankel, 2008; Rios *et al.*, 2019; Sundström and Oren, 2019; Yoshinaga, Nozaki and Wada, 2019) or dynamic glottal airway (Rosa *et al.*, 2003; Daily and Thomson, 2013; Jiang, Zheng and Xue, 2017). The typical Reynolds number of the glottal flow is about 2000 (Luo *et al.*, 2008). Considering the relative low Mach number (Mach<0.3) of the glottal flow (Scherer *et al.*, 2010), incompressible flow models have been most frequently employed. PDE based flow models are applied in

the studies of glottal flow deflection (Zheng, Mittal and Bielamowicz, 2011), intraglottal vortices (Mihaescu *et al.*, 2010; Oren, Khosla and Gutmark, 2014b) and the glottal pressure distribution (Scherer *et al.*, 2010; Jiang, Zheng and Xue, 2017, 2019).

Lighthill's acoustic analogy (Lighthill, 1982; Link *et al.*, 2009; Šidlof, Zörner and Hüppe, 2013, 2014) or the linear perturbed compressible equation for low Mach number condition (Seo and Moon, 2006; Seo and Mittal, 2011) can be utilized to obtain the acoustic solution based on the solution of the incompressible flow. Acoustic results can also be directly calculated through compressible flow models or slightly compressible models (Zhao, Frankel and Mongeau, 2001; Zhao *et al.*, 2002; Suh and Frankel, 2007; Larsson and Müller, 2009; Pickup and Thomson, 2011; Daily and Thomson, 2013; Farbos de Luzan *et al.*, 2015; Yoshinaga, Nozaki and Wada, 2019; Zhang, Krane and Yu, 2019). Currently, the compressible models have only been applied to two dimensional glottal flow.

1.3.3 Coupling

While assuming the glottal flow to be incompressible, the coupling calculation would involve three aspects: the solid tissue, the incompressible glottal flow and the acoustic (Alipour *et al.*, 2011) (Figure 1.5). In the two-way coupling of the solid tissue and the incompressible flow, the incompressible flow provides the loading force to drive the vocal fold motion and the deformed vocal fold provides the displacement and velocity boundary condition for the incompressible flow. For the calculation of the acoustic values, the linear source-filter theory assumes that the sound resonance in the vocal tract will not affect the vocal fold vibration. So the acoustic result calculated from the flow and tissue solution does not need to be fed back to the incompressible glottal flow or the vocal fold tissue. This allows more efficient calculation (Rosa *et al.*, 2003; Duncan, Zhai and Scherer, 2006; Tao *et al.*, 2007; Luo *et al.*, 2008; Link *et al.*, 2009; Mattheus and Brücker, 2011; Šidlof, Zörner and Hüppe, 2013).

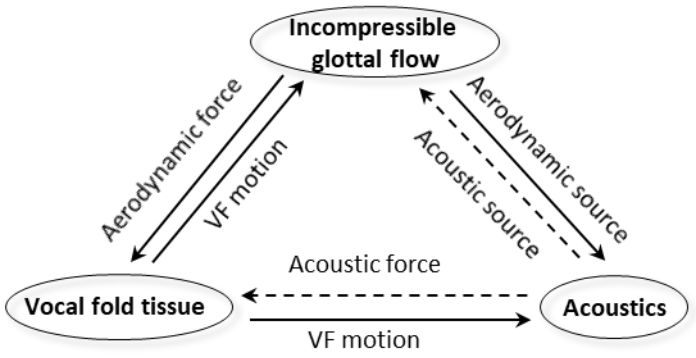


Figure 1.5 The fluid-structure-acoustic coupling.

In some conditions, however, the source-filter interaction is too strong to be ignored, for example, when the fundamental frequency is high and close to the formant of the glottal tract system which appears in children or singing (Titze, 2008; Migimatsu and Tokuda, 2019), or when the vocal folds have certain pathological conditions (Zhang, Neubauer and Berry, 2009). In these conditions, the acoustic resonance could result in register jump (Zhang, Neubauer and Berry, 2006b), changes in the fundamental frequency (Daily and Thomson, 2013), threshold pressure, flow pulse and further the vibration mode (Titze, 2008). One-way acoustic coupling is not able to capture these features because the contribution of the acoustic pressure to the vocal folds loading needs to be considered (Figure 1.5 dashed arrows)(Hatzikirou, Fitch and Herzel, 2006; Daily and Thomson, 2013; Maxfield, Palaparthi and Titze, 2016).

The moving boundary can be modeled using dynamic remeshing of the intraglottal fluid domain (Zhao *et al.*, 2002), the pseudo-solid approach (Link *et al.*, 2009) or the immersed-boundary method (Mittal *et al.*, 2008; Yang *et al.*, 2017). Among these approaches, the immersed-boundary method has

good capability in fluid-structure interaction applications especially for deformation of structures of complex geometries.

To deal with vocal folds contact, one common approach is the penalty force contact model where an external force is applied on vocal fold surfaces when they are getting very close to the contralateral side to prevent large penetration (Geng, Xue and Zheng, 2016; Zhang, 2019). In partial differential equation based flow solvers, a small artificial gap is required even in the closed phase to avoid the failure of the solver. For this purpose, collision planes are set close to the midsagittal plane to stop the vocal folds (Decker and Thomson, 2007; Luo, Mittal and Bielamowicz, 2009; Xue *et al.*, 2010; Pickup and Thomson, 2011; Daily and Thomson, 2013). However, doing so will introduce leakage flow. When the leakage flow is large, it might result in breathy voice and decrease in the maximum sound pressure level (Wang *et al.*, 2019).

1.4 Main research areas and findings

1.4.1 Vocal fold mechanics

The geometry and tension of the vocal folds play key role in the vocal fold dynamics and the acoustic outcome. These parameters vary between individuals and different phonation conditions. The causal relationship between physiological factors (tissue geometry, material property, muscle stimulation etc.) and the acoustic-related outputs (threshold pressure, pitch, intensity etc.) is not one-to-one correspondence or monotonic. Factors that might even have opposite effects are contributing together and carefully adjusted to achieve the final goal of the voice control. Main research areas include the threshold pressure, F_0 control, vocal register, vocal intensity and efficiency.

Threshold pressure is the minimum lung pressure required to initiate phonation (Titze, 1992a). A small value of the threshold pressure indicates an easy phonation onset. Typical value of the threshold

pressure for human larynx ranged in 0.13-0.34kPa (Chan, Titze and Titze, 1997; Berry, Chhetri and Neubauer, 2012). Generally, the threshold pressure can be decreased by decreasing the mucosal wave speed (Titze, 1988; Chan, Titze and Titze, 1997), decreasing the prephonatory gap (Titze, 1988; Titze, Schmidt and Titze, 1995; Zhang, 2009, 2010, 2016a; Perrine *et al.*, 2020), decreasing the vertical thickness of the medial surface (not applied to the very thin vocal fold) (Mendelsohn and Zhang, 2011; Zhang, 2016a), decreasing the stiffness in either layer of the vocal fold (Zhang, 2009, 2017a; Mendelsohn and Zhang, 2011), or changing to a more rectangular prephonatory glottal shape (Titze, 1988; Chan, Titze and Titze, 1997; Zhang, 2008; Wu and Zhang, 2019). Zhang (2010) explained the phonation onset through the eigen mode coupling of the first several modes. With linear stability analysis, the decrease of the prephonatory gap was found to have an effect of increasing the coupling strength of the eigen modes thus could decrease the threshold pressure. Geng, Xue and Zheng (2016) found that the vertical stiffness gradient also helps to decrease the threshold pressure in that the reduced stiffness at the superior aspect of the vocal fold allows a larger vibration. Moreover, when the vocal folds vibrate in acoustic driven style, featuring mainly in up-down motion instead of mucosal wave style motion, an inertive supraglottal tract or a compliant subglottal tract also helps to decrease the threshold pressure (Zhang, Neubauer and Berry, 2009). The favorable effect of an inertive supraglottal tract has also been reported in (Titze, 1988).

The control of the fundamental frequency is complex. A smaller vocal fold generally has higher fundamental frequency such as in women or children (Bowling *et al.*, 2017). With a certain vocal fold, main strategy to modulate fundamental frequency is adjusting the tension through CT/TA contraction. The increase of the tissue stiffness especially in the longitudinal direction significantly increases the fundamental frequency (Titze and Talkin, 1979; Zhang, 2009, 2016a, 2017a). The contraction of the CT

muscle that lengthens the vocal fold always increases the frequency. However, the effect of the contraction of the TA muscle depends (Titze, 1988; Titze, Jiang and Drucker, 1988; Chhetri and Park, 2016). If the vibration happens mainly in the cover layer, the fundamental frequency might decrease due to TA contraction since the cover layer slacks due to shortening. If the vibration reaches deep into the TA, then the fundamental frequency would increase (Titze, 1988; Titze, Jiang and Drucker, 1988). By analyzing vocal fold shape during singing over two octaves, Unteregger *et al.* (2018) suggested that in the three stages from F_0 to $2F_0$ and $4F_0$, the main muscles responsible for pitch control are the TA, CT and CT/TA together, respectively. In addition, vocal fold approximation (Zhang, 2016a), increased subglottal pressure (Titze and Talkin, 1979; Zhang, 2016a), the longitudinal gradient of the cross-section-area (Kelleher *et al.*, 2010) and the elastic modulus in the ligament (Bhattacharya, Kelleher and Siegmund, 2015) also contribute to higher fundamental frequency.

Vocal registers, characterized by different vocal fold vibration patterns, have distinct features on sound quality and the laryngeal energy transfer. The vibration pattern with a complete vocal fold closure has better excitation of the high-order harmonics. Factors contributing to the complete closure include large longitudinal stiffness, large medial surface thickness (Zhang, 2016a) and a stiff epithelium layer of the vocal fold (Xuan and Zhang, 2014; Tse, Zhang and Long, 2015). During speaking, the register transitions might happen are the modal-falsetto transition, which happens also in singing, and the pulse-modal transition, which might happen when pitch is low (~70Hz)(Titze, 2000). The modal register generally has strong mucosal-wave propagation, complete vocal fold closure and shadow spectral slope (Titze, 2000). The falsetto register has much weaker mucosal wave propagation, vocal fold closure and steep spectral slope. The relative contraction strength of the CT and TA muscles influences the modal-falsetto transition (Hirano, 1974; Story and Titze, 1995; Zhang, 2009). When TA contracts more than CT

or both in low level, the vocal fold is thick (bulging out especially in the inferior part) and soft in the cover layer, enabling the mucosal-wave propagation in the entire vocal fold, resulting in modal register. Contrarily, when CT contracts much more powerfully than TA, the vocal fold would be thin (small vertical thickness) with high tension in all the layers. It results in weak mucosal-wave propagation and the falsetto register (Hirano, 1974; Tokuda *et al.*, 2007). Zhang (2009, 2017a) suggested that the change of the vertical thickness of the medial surface rather than the stiffness resulted from the TA/CT activation might be the main reason of the modal-falsetto transition. The fact that these two are usually affected simultaneously makes it difficult to differentiate the cause. The glottal tract system also influences register transition (Titze, 2000). In the acoustically driven type of vibration that features strong up-down motion, the fundamental frequency is mainly decided by the formant of the glottal tract (Zhang, Neubauer and Berry, 2006a, 2006b, 2009; Daily and Thomson, 2013; Migimatsu and Tokuda, 2019). Daily and Thomson (2013) showed that when the length of the subglottal tract changed from 50-100cm to 250-300cm, it changed from 1/4 to 3/4 wavelength resonator, resulting in frequency jump.

Vocal intensity is proportional with subglottal pressure (Titze and Talkin, 1979; Zhang, 2016a) and vocal fold length (Titze and Talkin, 1979; Zhang, 2016a). The vertical stiffness gradient (Geng, Xue and Zheng, 2016) also has positive effect on it and it is more prominent at low subglottal pressures.

Vocal efficiency is defined as the ratio of radiated sound power to aerodynamic power (Titze, 1992b; Titze, Maxfield and Palaparthi, 2016). Mass lesion of the vocal folds would significantly reduce the efficiency (Jiang *et al.*, 2004). Strategies to improve the efficiency include the bulging of the medial surface (Titze and Talkin, 1979), the increase of the body-cover stiffness ratio (adjusted by TA/CT muscles)(Zhang, 2009) and the increase of the longitudinal stiffness especially in the cover layer (Zhang, 2014). The maintenance of the vocal fold adduction through LCA/IA contraction (Chhetri and Park, 2016)

and the excitation of high-order modes that have more lateral motion(Zhang, 2009) also increase the vocal efficiency. Moreover, vocal efficiency was observed to be higher in adult than in children, suggesting a favorable effect from the maturation of the vocal ligament (Tang and Stathopoulos, 1995). Titze, Maxfield and Palaparthi (2016) suggested that a pressure conversion ratio might serve as the alternate of the efficiency due to the difficulty of the combined measurement of lung pressure and tracheal airflow.

1.4.2 Flow field

In the glottis, flow separation occurs slightly downstream of the minimum area location when the vocal folds are closing in a divergent shape. When Reynolds number is below 2000, flow separation might happen when the divergent angle is just 5 degrees (Sparrow, Abraham and Minkowycz, 2009). Flow separation affects the vibration pattern by influencing the pressure distribution on the vocal fold surface and introduces vortex structures that will contribute to the quadrupole sound source. When the wall shear force is zero, it indicates the occurrence of the flow separation (Pelorson *et al.*, 1994; Cisonni *et al.*, 2010). The flow separation location can also be determined manually from streamline through experiment visualization (Cisonni *et al.*, 2010; Šidlof *et al.*, 2011; Stewart, Erath and Plesniak, 2014) or numerical simulation (Smith and Thomson, 2012). The cross-section area ratio of where flow separates over the minimum area is found to be around 1~2 (Pelorson *et al.*, 1995; Alipour and Scherer, 2004; Sciamarella and Le Quéré, 2008; Smith and Thomson, 2012) but can be much larger (5.5-8) shortly prior to the final close of the vocal fold (Decker and Thomson, 2007; Šidlof *et al.*, 2011). Upstream movement of the flow separation within a divergent glottis was found to facilitate phonation (Zhang, 2008). Several factors might result in the upstream movement, such as a more divergent shape (Alipour and Scherer, 2004; Li *et al.*, 2006; Zhang, 2008), larger flow (Alipour and Scherer, 2004; Li *et al.*, 2006; Sciamarella and

Le Quéré, 2008) and the outward motion of the vocal folds (Alipour and Scherer, 2004; Sciamarella and Le Quéré, 2008).

Flow separation results in intraglottal vorticities. Vortex structures generated by the glottal jet shear layer instability travel downstream with the glottal jet at the speed about 50%-70% (Zhang, Mongeau and Frankel, 2002; Oren, Khosla and Gutmark, 2014a) of the jet speed. PIV measurements (Oren, Khosla and Gutmark, 2014a) and numerical studies (Mihaescu *et al.*, 2010; Farbos de Luzan *et al.*, 2015) suggested that the intraglottal vortex structures during vocal fold closing had larger intensity under higher subglottal pressure condition. Such vortex structures that could result in negative pressure as high as about 55% of the value of the subglottal pressure (Oren, Khosla and Gutmark, 2014b) were supposed to have the effect of facilitating vocal fold closing and increasing the glottal waveform skewness, maximum flow deceleration rate and sound pressure level. However, by applying an external force with the similar value as provided in Oren, Khosla and Gutmark (2014b) downstream the flow separation location to mimic the effect of the intraglottal vortices, Farahani and Zhang (2014) found that the intraglottal vortices could only increase the MFDR by 12.5% corresponding to 1 dB increase in sound pressure level. By directly introducing the vortex structure near a plate, Pirnia *et al.* (2018) quantified that the effect of the intraglottal vortices on vocal fold dynamics might be negligible. Therefore, the effect of the intraglottal vortices on vocal fold dynamics is still under debate.

Jet deflection, i.e. glottal jet attaching to one of the vocal folds walls, which may happen even in a symmetric glottis could result in pressure difference of the two walls to about 5-6% subglottal pressure value (Scherer *et al.*, 2001). It influences the dipole sound source (Erath and Plesniak, 2006b) and could be a factor of normal jitter values (Scherer *et al.*, 2001). Numerically, a partial-differential equation based fluid solver is needed to capture this phenomenon (Tao *et al.*, 2007). One possible

mechanism of the jet deflection is known as the Coanda effect (Erath and Plesniak, 2006c) that the flow has the tendency to adhere to the adjacent surface. Experiment measurement on static vocal fold models showed that when the divergent angles were 10 and 20 degrees, the glottal flow attached to one side of the vocal folds. But this did not happened in the 40 degrees case (Erath and Plesniak, 2006a, 2006c). Such jet deflection behavior was explained by referring to the maps of diffuser flow regimes: contrary to 10-20 degrees that are in the unstable transition regime, 40 degrees is in the fully developed stall regime, thus there is no flow attachment (Erath and Plesniak, 2006a). However, Pelorson *et al.* (1994) suggested that Coanda effect might not have enough time to establish itself in such small time range of vocal fold vibration. Another possible mechanism of jet deflection proposed by Zheng *et al.* (2011) and Zheng, Mittal and Bielamowicz (2011) is that it might be resulted from the downstream asymmetric vortex structures. They observed from the numerical results that the jet deflection was always preceded by large-scale asymmetry in the downstream portion of the glottal jet. Three-dimensional jet has much smaller jet deflection compared with two-dimensional jet (Mattheus and Brücker, 2011; Zheng *et al.*, 2011; Xue *et al.*, 2014) and is even more stable in realistic configuration compared with simplified shape (Xue *et al.*, 2014). The overestimation of the jet deflection in the two dimensional model might artificially generate low-frequency pattern in the acoustic spectrum due to the flow recirculation in the supraglottal space (Mattheus and Brücker, 2011).

Features unique to the three-dimensional glottal flow include the jet axis switching and the longitudinal asymmetry. Axis switching of the jet is the phenomenon that the glottal jet narrows in the longitudinal direction but widens in the lateral direction when propagating downstream of the vocal folds (Khosla *et al.*, 2008; Triep and Brücker, 2010; Zheng *et al.*, 2011; Krebs *et al.*, 2012; Kirmse and Brücker, 2014). Vortices resulted from the axis switching might affect voice quality (Khosla *et al.*, 2008).

For the longitudinal asymmetry, Alipour and Scherer (2000) found that the pressure drop near the mid-coronal plane was larger than the two ends. Pressure near two ends might be positive in the whole cycle which protects the vocal folds during collision. Such longitudinal pressure gradient might be more prominent when the glottis is a divergent shape. Secondary flow resulted from it can be around 10% of the axial velocity (Scherer *et al.*, 2010). Khosla *et al.* (2008) observed that the glottal jet skewed toward the anterior end above the folds. From numerical study of a subject-specific vocal fold model, Xue *et al.* (2014) observed that both subglottal pressure and subglottal angle increased from the anterior to the posterior end. This resulted in a greater vertical motion at the posterior end which delayed the glottal opening.

The effect of the ventricular folds on the glottal dynamics is another focus in the flow field. Ventricular folds, also knowns as the false vocal folds (FVFs), serve primarily as a protecting valve during swallowing (Kniesburges *et al.*, 2017) and barely move in normal phonation. FVFs move in some special conditions such as glottal stops (Birk *et al.*, 2016) or some special singing conditions (Kniesburges *et al.*, 2017). In FVF-related dysphonia, FVFs disturb phonation, resulting in hoarseness and change in pitch (Jiao *et al.*, 2018). Stationary FVFs could straighten the glottal jet, i.e., suppress the jet deflection (Kucinschi *et al.*, 2006; Bailly *et al.*, 2008; Triep and Brücker, 2010; Kirmse and Brücker, 2014; Farbos de Luzan *et al.*, 2015; Xue and Zheng, 2017; Sadeghi *et al.*, 2018). Contrarily, the vibration of the FVFs decreased voice quality (Birk *et al.*, 2016). When the FVF motion is in different phase with the vocal folds (Bailly, Henrich and Pelorson, 2010; Alipour and Scherer, 2012), it could result in irregular vibration pattern (Moisik and Esling, 2014).

Many studies reported that the FVF decreased the threshold pressure (Zheng *et al.*, 2009; Birk *et al.*, 2016; Kniesburges *et al.*, 2017) by decreasing the pressure level in the ventricle (Alipour and Scherer,

2012; Kniesburges *et al.*, 2017). The minimum ventricle pressure might be achieved when the FVF-VF gap ratio is 2-5.3 (Agarwal, 2004; Farbos de Luzan *et al.*, 2015; Bailly *et al.*, 2008). Kniesburges *et al.* (2017) and Sadeghi, Döllinger, *et al.* (2019) found that the FVF has the effect of increasing the glottal flow resistance. Contrarily, the laryngeal flow resistance might decrease (Zhang *et al.*, 2002; Zheng *et al.*, 2009; Farbos de Luzan *et al.*, 2015; Sadeghi, Döllinger, *et al.*, 2019). Xue and Zheng (2017) suggested that the final effect of FVF on laryngeal flow resistance might be the combined effect of the reduced mixing-related loss and the increased friction-related loss, both of which were highly related to the FVF gap.

The FVF-vocal fold gap ratio could also affect the fundamental frequency. When FVF-VF gap ratio increased from 0 to 1, F_0 would increase and keep around a constant value when the ratio further increased (Bailly *et al.*, 2008). Birk *et al.* (2016) suggested that the FVFs wound confine the F0 increase. FVFs were also reported to have the effect of increasing the glottal flow rate (Sadeghi *et al.*, 2018; Sadeghi, Kniesburges, *et al.*, 2019), increasing the sound intensity (Alipour, Jaiswal and Finnegan, 2007), enhancing the monopole sound source (Zheng *et al.*, 2009) and generating stronger glottal vortices (Farbos de Luzan *et al.*, 2015; Sadeghi, Kniesburges, *et al.*, 2019). McGowan and Howe(2010) reported that the effect of the FVFs on the output acoustic pressure was small when the movement of the vocal folds was prescribed. However, Farahani *et al.* (2013) suggested a back-coupling interaction between vocal folds and FVFs that the shedding vortex might rebound from FVFs and reach vocal folds again. Such cavity oscillation could generate high-order harmonic component in the acoustic output as a dipole source (Zhang *et al.*, 2002). Alipour, Jaiswal and Finnegan (2007) suggested that FVF might be responsible for the low-frequency components in the acoustics and for oscillation instabilities.

1.4.3 Acoustics

Analysis with acoustic analogy (Ffowcs Williams and Hawkings, 1969) shows that major sources contributing to the sound generation in the larynx are (1) a dipole source due to the unsteady axial forces exerted by the walls from the fluid; (2) a monopole source due to the mass of the volume of air displaced by the motion of the orifice walls; and (3) a quadrupole source related to kinetic energy fluctuations of the flow downstream of the orifice (Zhao *et al.*, 2002). Dipole source is the principal sound source. McPhail, Campo and Krane (2019) demonstrated it by showing a strong correlation between the transglottal pressure and the radiated sound. In the 125Hz frequency condition, the strength of the monopole source is just around 1/5 of the dipole source. Dipole source from vortex shedding and the interaction with the vocal folds walls is responsible for the tonal sound (Suh and Frankel, 2007; Khosla *et al.*, 2008). Moreover, Zhang and Mongeau (2006) suggested that the dipole source also included a random component caused by turbulent fluctuating pressures acting on the superior vocal folds walls. This random component and the quadrupole source together contribute to the broadband sound, which is dominant in high frequency and unvoiced sound (Zhang and Mongeau, 2006).

Glottal flow rate affects the sound source by influencing the jet pattern. When the jet core is short such as in the low flow rate condition, the vortex structures interact with the vocal folds walls in the glottis. It contributes to the dipole source. However, when the jet presents to be a long transition one in the high flow rate condition, the vortex structures would generate away from any surfaces, contributing to the quadrupole source (Kucinschi *et al.*, 2006). FVFs affect the dipole source by stabilizing and straightening the glottal jet (Kucinschi *et al.*, 2006) or directly interacting with the jet (Erath and Plesniak, 2006b). Frequency also influences the sound source. Since the strength of the

monopole source is proportional with the frequency while the dipole is not, the monopole source can be significant when the frequency increase to a high level, such as 400Hz as suggested by Zhao *et al.* (2002). Increased frequency also brings about more small-scale vortex structures which contribute to the broadband sound (Zhang and Mongeau, 2006). Sundström and Oren (2019) showed that in the nasal cavity, turbulence interacting with the nasal cavity walls generated the dipole source that was two orders of magnitude larger than the quadrupole source that generated by turbulence itself.

1.5 Objectives of the current study

The capability and complexity of computational modelling of phonation have significantly improved in the recent decades in both fluid and tissue aspects. However, challenged by the computational cost, most of the flow models are still based on Bernoulli flow and the vocal fold models are based on finite element method but with simplified geometry. The Bernoulli flow based models are much quicker and easier compared with the PDE based flow models but are limited since they are linear, and the viscous effects are neglected. Moreover, by applying uniform pressure in each transverse plane, the pressure loading on vocal folds is inaccurate especially when the airway shape is complex. Contrarily, three dimensional PDE based flow solution would describe the flow domain more accurately thus help to improve the understanding of the glottal flow dynamics and the FSI process.

Another limitation of current studies is that the numerical models are usually validated by comparing with physiological possible range, without the verification against the specific object. Difficulties in the validation include the coupled modelling of the three-dimensional vocal fold vibration and the glottal flow and the uncertainly of material properties. Though there are literatures reporting the measurements of the material properties of the tissue, the data vary in a large range between individuals. Ideally, when performing the validation, the measurement of vocal fold vibration, the

geometry and the material property should come from the same subject. By applying the geometry and material property obtained from the experiment measurement in the simulation setup, a one-to-one comparison could be performed.

Given the above limitations of the current studies, the objectives of the current study are (1) performing more rigorous validation of the numerical approach by performing one-to-one comparison of the numerical results with the experiment measurements, and (2) performing FSI simulations of the laryngeal flow with PDE based flow model to extend our current understanding about the source of the voiced sound, especially the effect from the layered structure of the vocal fold. The dissertation is outlined as follows:

Chapter 2 describes the computational method that explicitly couples a sharp-interface immersed boundary method based incompressible fluid solver, a hydrodynamic/acoustics splitting method based acoustics solver and a viscoelastic solid solver.

Chapter 3 provides the validation of the numerical solver against the experiment measurements on a canine larynx and pigeon syringes.

Chapter 4 provides the application on a simplified human vocal fold and vocal tract system, which would serve as the baseline case of the parametric study in later chapters. The vocal fold has been considered as a three-layer structure including the cover, the ligament, and the body layers. In Chapter 5 and 6, the degree of the longitudinal variation of the thickness of the cover and ligament layers, and the cover-body layer thickness ratio have been systematically varied in a large range. The effect on flow rate related parameters, vocal fold dynamics and the glottal energy transfer has been investigated. Chapter 7 discusses the conclusions, limitations, and possible future works.

CHAPTER 2 NUMERICAL METHOD

In the current study, a high-fidelity numerical solver that couples a sharp-interface immersed boundary method based incompressible fluid solver, a hydrodynamic/acoustics splitting method based acoustics solver and a viscoelastic solid solver was employed. This chapter provides the brief description of the numerical solvers. More details of the solvers could be referred to (Luo *et al.*, 2008; Mittal *et al.*, 2008; Zheng, 2009; Seo and Mittal, 2011).

2.1 The fluid solver

The fluid solver is based on the immersed boundary method which was firstly proposed by (Peskin, 1972) in the cardiac application. By solving the governing equation on fixed Cartesian grid, it avoids the need to regenerate the grid during fluid-structure interaction and has the advantage in describing the complex tissue boundary in the vocal fold application. Two sets of grids are included in the fluid solver: (1) the Cartesian grid, where the governing equation of the fluid is solved and (2) the immersed boundary representing the shape of the fluid region. In the current phonation application, the immersed boundary presents the airway. Through fluid-structure coupling, the information about the tissue displacement/velocity would be transferred to the immersed boundary to build the boundary condition of the fluid domain.

2.1.1 Governing equation and numerical scheme of the incompressible flow

The Mach number in the glottal flow is less than 0.3 (Scherer *et al.*, 2010). Therefore, the incompressible simplification, which has been employed in most of the studies as stated in Chapter 1, has also been applied in the flow field in the current study. The governing equation of the fluid is the three-dimensional, unsteady, viscous, incompressible Navier-Stokes equations:

$$\nabla \cdot \vec{U} = 0 \tag{2.1}$$

$$\frac{\partial \vec{U}}{\partial t} + \left(\vec{U} \cdot \nabla\right)\vec{U} = -\frac{1}{\rho_0}\nabla p + \upsilon_0 \nabla^2 \vec{U} \tag{2.2}$$

where $\vec{U}$, ρ_0, p, υ_0 are the incompressible flow velocity, density, pressure, and kinematic viscosity, respectively. The governing equation is solved in Cartesian grids and the primitive variables ($\vec{U}$, p) are stored at the cell center. Time marching of (2.2 is based on fractional-step method (van Kan, 1986). The equation is split as:

$$\frac{u^* - u^n}{\Delta t} = -\frac{1}{\rho_0}\nabla p^n + \upsilon_0 \nabla^2 \vec{U} - \left(\vec{U} \cdot \nabla\right)\vec{U} \tag{2.3}$$

$$\frac{u^{n+1} - u^*}{\Delta t} = -\frac{1}{\rho_0}\nabla p' \tag{2.4}$$

where u^* is an intermediate velocity, p' is pressure correction variable. Sub-steps have been taken to obtain the value of the variables (Mittal *et al.*, 2008):

Step 1. (2.3 is calculated at the cell nodes to obtain u_i^*:

$$\frac{u_i^* - u_i^n}{\Delta t} + \frac{1}{2}(3N_i^n - N_i^{n-1}) = -\frac{1}{\rho_0}\frac{\delta p^n}{\delta x_i} + \frac{1}{2}(D_i^* + D_i^n) \tag{2.5}$$

where $N_i = \frac{\delta(U_j u_i)}{\delta x_j}$ is the convective terms employing the Adams-Bashforth scheme with U_j representing the face-center velocities, $D_i = \nu \frac{\delta}{\delta x_j}(\frac{\delta u_i}{\delta x_j})$ is the diffusive term employing the implicit Crank-Nicolson scheme, $\frac{\delta}{\delta x}$ denotes a second-order central difference scheme. (2.5 is solved using a line-SOR scheme.

Step 2. The divergence of (2.4 generates the Poisson equation to obtain the pressure correction variable p' as u^{n+1} is divergence free according to the mass continuity equation ((2.1):

$$\frac{1}{\rho_0}\frac{\delta}{\delta x_i}\left(\frac{\delta p'}{\delta x_i}\right) = \frac{1}{\Delta t}\frac{\delta U_i^*}{\delta x_i} \tag{2.6}$$

The pressure Poisson equation is solved through a highly efficient multi-grid method that employs a Gauss-Siedel line-SOR smoother (Dong, Mittal and Najjar, 2006).

Step 3. The pressure and velocity values are corrected using p':

$$p^{n+1} = p^n + p' \tag{2.7}$$

$$u_i^{n+1} = u_i^* - \frac{\Delta t}{\rho_0}\left(\frac{\delta p'}{\delta x_i}\right)_{cc} \tag{2.8}$$

$$U_i^{n+1} = U_i^* - \frac{\Delta t}{\rho_0}\left(\frac{\delta p'}{\delta x_i}\right)_{fc} \tag{2.9}$$

where *cc* and *fc* denote cell-center and face-center, respectively.

2.1.2 Treatment of boundary condition

The deformation of the immersed boundary provides the boundary condition for the incompressible flow calculation. Sharp-interface immersed boundary method has been employed to impose the boundary condition for the fluid domain (Mittal *et al.*, 2008).

Based on the location of the immersed boundary, the Cartesian grids in the computational domain would first be differentiated as solid or fluid. Ghost cell (GC) is defined as the solid cell that has at least one fluid cell neighbor. In the ghost-cell methodology, a probe that perpendicularly intersects with the immersed boundary is generated from the GC to the image point (IP) (Figure 2.1(a)). The boundary intercept is named as BI. Values of the fluid nodes surrounding IP are interpolated to obtain the fluid value at IP:

$$\Phi_{IP} = \sum \beta_i \Phi_i \tag{2.10}$$

where Φ_{IP} is the fluid value at IP, Φ_i is the fluid value at the nodes surrounding the IP, β_i is the interpolating coefficient, i ranges from 1-4 in the 2D problem and 1-8 in the 3D problem. For Dirichlet

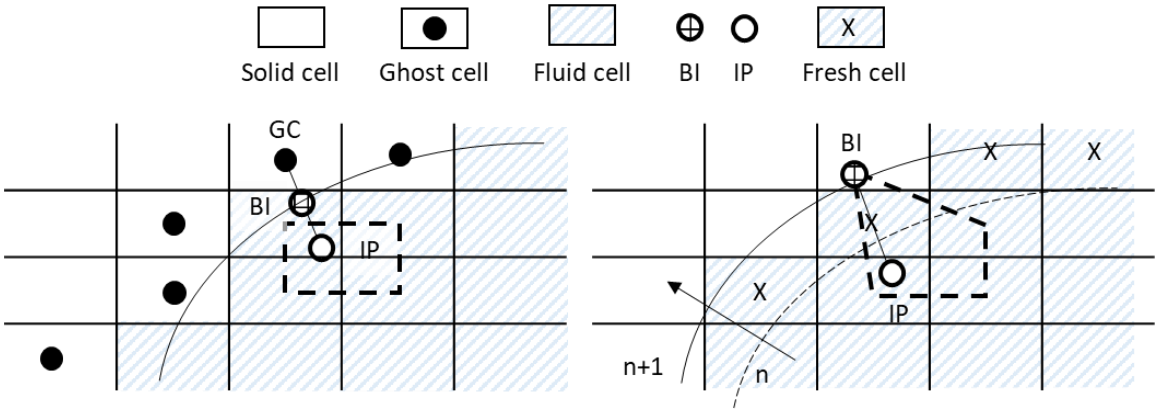
boundary condition, value at GC is obtained by linear interpolation as shown in (2.11 when BI locates in the middle of GC and IP:

$$\Phi_{GC} = 2\Phi_{BI} - \Phi_{IP} \tag{2.11}$$

where Φ_{GC} is the fluid value at GC, Φ_{BI} is the value at BI according to the displacement and velocity of the immersed boundary. For Neumann boundary condition,

$$\Phi_{GC} = \Phi_{IP} + \Delta l(\frac{\delta \Phi}{\delta n})_{BI} \tag{2.12}$$

Where *Δl* is the length of probe. In this boundary formulation methodology, second-order accuracy of the velocity variables can be achieved locally and globally (Mittal *et al.*, 2008).



(a) (b)

Figure 2.1 (a) Ghost cell method; (b) The treatment of fresh cell.

In the condition of a moving boundary, whether the cell is a fluid cell or a solid cell depends on the immersed boundary location at the specific time instant. Two types of special cells are termed as "Fresh cell" or "Dead cell". "Fresh cell" refers to the cell that was solid in the previous time step but becomes fluid in the current time step. "Dead cell" refers to the cell that was fluid in the previous time step but becomes solid in the current time step. Special treatment is required to obtain the value of the

fresh cell. As shown in Figure 2.1(b), BI is generated from probe pointing from fresh cell perpendicularly to the immersed boundary. Fresh cell locates in the middle of the BI and IP. Φ_{IP} is interpolated using surrounding nodes and BI: $\Phi_{IP} = \Phi_{BI} + \sum \beta_i \Phi_i$ where i ranges 1-3 in the 2D case and 1-7 in the 3D case in the surrounding nodes excluding the fresh cell. Then the fresh cell value (Φ_{Fresh}) is obtained as: $\Phi_{Fresh} = (\Phi_{BI} + \Phi_{IP})/2$ in the case that it locates in the middle between BI and IP.

During dynamic case, at the beginning of each time step, the immersed boundary is deformed according to the tissue motion. Then the GC values and the IP values are updated accordingly. The governing equations is iterated on the Cartesian grid until the convergence criteria has been satisfied.

2.1.3 Acoustic solver

In the glottal flow circumstance, direct numerical simulation of such low Mach number aeroacoustics faces the challenge of scale disparities between the hydrodynamic vortical motions and the acoustic waves (Seo and Moon, 2006). Alternatively, a hydrodynamic/acoustics splitting method based acoustics solver which is more computational efficient has been integrated with the incompressible fluid solver (Seo and Mittal, 2011). The method has been validated against the fundamental dipole/quadrupole noise problem and the low Mach turbulent flow noise problems (Seo and Moon, 2006, 2007; Moon *et al.*, 2010). In the hydrodynamic/acoustics splitting method, the total flow variables are decomposed into the incompressible variables and the perturbed compressible ones:

$$\begin{aligned} \rho(\vec{x},t) &= \rho_0 + \rho'(\vec{x},t) \\ \vec{U}(\vec{x},t) &= \vec{u}(\vec{x},t) + \vec{u}'(\vec{x},t) \\ P(\vec{x},t) &= p(\vec{x},t) + p'(\vec{x},t) \end{aligned} \tag{2.13}$$

where ρ, $\vec{U}$, P are the total flow density, velocity and pressure, respectively, ρ_0, $\vec{u}$, p are incompressible perturbed flow density, velocity, pressure, respectively, and ρ', $\vec{u}'$, p' are compressible perturbed flow

density, velocity, pressure, respectively. The calculation of the incompressible variables has been discussed before. The perturbed compressible parts employ the linearized perturbed compressible equation (LPCE) (Seo and Moon, 2006; Seo and Mittal, 2011):

$$\frac{\partial \rho'}{\partial t} + (\vec{u} \cdot \nabla)\rho' + \rho_0(\nabla \cdot \vec{u}') = 0$$

$$\frac{\partial \vec{u}'}{\partial t} + \nabla(\vec{u}' \cdot \vec{u}) + \frac{1}{\rho_0}\nabla p' = 0 \qquad (2.14)$$

$$\frac{\partial p'}{\partial t} + (\vec{u} \cdot \nabla)p' + \gamma p(\nabla \cdot \vec{u}') + (\vec{u}' \cdot \nabla)p = -\frac{Dp}{Dt}$$

where γ is the ratio of the specific heat. The $\frac{Dp}{Dt}$ term on the RHS represents the sound source from the flow solver. The LHS presents the effects of acoustic wave propagation and refraction in the unsteady, inhomogeneous base flow. The boundary condition on the solid wall is:

$$\frac{\partial \rho'}{\partial n} = 0, \qquad \frac{\partial p'}{\partial n} = 0, \qquad \vec{u}' \cdot \hat{n} = 0$$

where $\hat{n}$ is the unit vector of face normal direction. The LPCE is discretized with a sixth-order central compact finite difference scheme in space and integrated using a four-stage Runge-Kutta method in time. The LPCE based acoustic solver has been coupled with the IBM based incompressible fluid solver. Further details of this model can be found in (Seo and Moon, 2006; Seo and Mittal, 2011).

2.2 The solid solver

A finite element based three-dimensional solid solver is employed to solve the tissue dynamics (vocal fold vibration). Details has been discussed below.

2.2.1 Constitutive law

Based on the discussion in Chapter 1, vocal fold vibration around a given postured position is generally considered as small deformation with much larger deformation in the transverse plane than in

the longitudinal direction. Therefore, most of the studies studying the vocal fold vibration applies the linear assumption and many of them also employed the transversely isotropic assumption (Titze and Talkin, 1979; Berry *et al.*, 1994; Cook and Mongeau, 2007; Kelleher *et al.*, 2010; Xue *et al.*, 2014; Zhang, 2016a, 2017a; Erath, Zañartu and Peterson, 2017). Following these, in the current study, the material property of the tissue is assumed to be incompressible transversely isotropic. The Kelvin-Voigt model has been employed for modeling the viscoelastic material (Fung, 1993):

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\varepsilon} + \boldsymbol{\eta}\dot{\boldsymbol{\varepsilon}} \tag{2.15}$$

where $\boldsymbol{\sigma}$ is the stress tensor, $\boldsymbol{\varepsilon}$ is the strain tensor, $\dot{\boldsymbol{\varepsilon}}$ is the strain rate tensor, $\mathbf{C}$ and $\boldsymbol{\eta}$ are the material constant tensors corresponding to elasticity and viscosity, respectively. In the condition that x-y plane is the transverse plane and z is the longitudinal direction, the strain-stress relationship takes the form:

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \tau_{yz}\\ \tau_{zx}\\ \tau_{xy}\end{bmatrix} = \begin{bmatrix} \frac{1-\nu_{pz}\nu_{zp}}{E_p E_{pz}\Delta} & \frac{\nu_p+\nu_{pz}\nu_{zp}}{E_p E_{pz}\Delta} & \frac{\nu_{zp}+\nu_p\nu_{zp}}{E_p E_{pz}\Delta} & 0 & 0 & 0 \\ \frac{\nu_p+\nu_{pz}\nu_{zp}}{E_p E_{pz}\Delta} & \frac{1-\nu_{pz}\nu_{zp}}{E_p E_{pz}\Delta} & \frac{\nu_{zp}+\nu_p\nu_{zp}}{E_p E_{pz}\Delta} & 0 & 0 & 0 \\ \frac{\nu_{pz}+\nu_p\nu_{pz}}{{E_p}^2\Delta} & \frac{\nu_{pz}(1+\nu_p)}{{E_p}^2\Delta} & \frac{1-{\nu_p}^2}{{E_p}^2\Delta} & 0 & 0 & 0 \\ 0 & 0 & 0 & G_{pz} & 0 & 0 \\ 0 & 0 & 0 & 0 & G_{pz} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{E_p}{2(1+\nu_p)} \end{bmatrix} \begin{bmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \varepsilon_{zz}\\ \gamma_{yz}\\ \gamma_{zx}\\ \gamma_{xy}\end{bmatrix} \tag{2.16}$$

where $\Delta = \frac{(1+\nu_p)(1-\nu_p-2\nu_{pz}\nu_{zp})}{{E_p}^2 E_{pz}}$. τ_{yz}, τ_{zx}, τ_{xy} are the shear stresses in y-z, z-x, x-y planes, respectively; γ_{yz}, γ_{zx}, γ_{xy} are the engineering shear strains in y-z, z-x, x-y planes equaling to $2\varepsilon_{yz}$, $2\varepsilon_{zx}$, $2\varepsilon_{xy}$, respectively. E_p, E_{pz} stands for transverse and longitudinal Young's modulus, respectively; G_{pz} is the shear modulus in longitudinal direction; ν_p, ν_{pz} are the transverse and longitudinal Poisson's ratio, respectively. The symmetry of the stiffness matrix requires that $\frac{\nu_{pz}+\nu_p\nu_{pz}}{{E_p}^2\Delta} = \frac{\nu_{zp}+\nu_p\nu_{zp}}{E_p E_{pz}\Delta}$. Thus ν_{zp} can be calculated as $\nu_{zp} = \nu_{pz}E_{pz}/E_p$.

The strain rate-stress relationship takes the form:

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \tau_{yz}\\ \tau_{zx}\\ \tau_{xy}\end{bmatrix} = \begin{bmatrix}\eta & 0 & 0 & 0 & 0 & 0\\ 0 & \eta & 0 & 0 & 0 & 0\\ 0 & 0 & \eta & 0 & 0 & 0\\ 0 & 0 & 0 & \eta/2 & 0 & 0\\ 0 & 0 & 0 & 0 & \eta/2 & 0\\ 0 & 0 & 0 & 0 & 0 & \eta/2\end{bmatrix}\begin{bmatrix}\dot{\varepsilon_{xx}}\\ \dot{\varepsilon_{yy}}\\ \dot{\varepsilon_{zz}}\\ \dot{\gamma_{yz}}\\ \dot{\gamma_{zx}}\\ \dot{\gamma_{xy}}\end{bmatrix} \tag{2.17}$$

where η is the viscosity of the tissue. It is assumed that the strain rate-stress relationship is not affected by the Poisson effect.

2.2.2 Governing equation and numerical scheme

The governing equation of the tissue dynamics is the Navier equation:

$$\rho_{tiss}\frac{\partial^2 d_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} + \rho_{tiss} f_i \tag{2.18}$$

where *i* and *j* range from 1-3, ρ_{tiss} is the tissue density, d_i is the displacement, σ_{ij} is the stress tensor, and f_i is the body force. By employing the Galerkin method as the finite element formulation, (2.18 becomes:

$$M_{\alpha\beta}\ddot{d}^{\beta} + C_{\alpha\beta}\dot{d}^{\beta} + K_{\alpha\beta}d^{\beta} = F_{\alpha}^{t} + F_{\alpha}^{b} \tag{2.19}$$

where $\ddot{d}$ and $\dot{d}$ are the second and first derivative of nodal displacement, $M_{\alpha\beta} = \int_v \rho_{tiss} N_{j\beta} N_{i\alpha} dv$ is the mass matrix, $C_{\alpha\beta} = \int_v A_{ijkl} B_{kl\beta} B_{ij\alpha} dv$ is the damping matrix, $K_{\alpha\beta} = \int_v C_{ijkl} B_{kl\beta} B_{ij\alpha} dv$ is the stiffness matrix, $F_{\alpha}^{t} = \int_{s\sigma} \sigma_s N_{i\alpha} ds$ is the traction force on node α, $F_{\alpha}^{b} = \int_{s\sigma} \rho_{tiss} f_i N_{i\alpha} dv$ is the body force on node α and *N* is the shape function. C_{ijkl} and A_{ijkl} are the tensor forms of the material coefficients in Equation 2.16 and Equation 2.17, respectively.

Discretizing (2.19 in time using a second-order Newmark scheme results in the discretized equation as shown in (2.20:

$$\left(K+\frac{1}{\beta\Delta t^2}M+\frac{\gamma}{\beta\Delta t}C\right)D^{n+1}$$

$$=F^{n+1}+M\left[\frac{1}{\beta\Delta t^2}D^n+\frac{1}{\beta\Delta t}\dot{D}^n+\left(\frac{1}{2\beta}-1\right)\ddot{D}^n\right] \tag{2.20}$$

$$+C\left[\frac{\gamma}{\beta\Delta t}D^n+(\frac{\gamma}{\beta}-1)\dot{D}^n+\left(\frac{\gamma}{2\beta}-1\right)\Delta t\ddot{D}^n\right]$$

where β and γ are constants equaling to 0.25 and 0.5, respectively, in the current study. This provides second-order accuracy. The equation is solved by a banded LU decomposition solver and the Cutill-Mckee and Gibbs-Poole-Stockmeyer methods are used to re-index the nodes to produce a banded matrix. Details regarding the numerical algorithm of the flow and solid solvers can be found in (Zheng, 2009; Zheng *et al.*, 2010; Xue, 2011).

2.2.3 Contact model

Vocal fold vibration will result in contact of the two sides vocal folds. Considering that the cases we would study employ the symmetric or near-symmetric vocal fold models, two contact planes have been set at the locations of $x\pm\Delta x$ for the corresponding vocal folds, where x denotes the lateral location of the mid-sagittal plane and Δx denotes the offset of the contact plane. When the vocal fold vibrates to the position exceeding the contact plane, it is forced to stop at the contact plane and both velocity and acceleration are set to zero. An artificial gap ($2\Delta x$) would be maintained between the two vocal folds for the success of the fluid solver, which is also common in other partial-differential equation based fluid solvers (Suh and Frankel, 2007; Luo, Mittal and Bielamowicz, 2009; Daily and Thomson, 2013; Sadeghi, Kniesburges, *et al.*, 2019). Though necessary, such artificial gap introduces leakage flow even when the vocal fold has been closed, which might result in breathy sound.

2.3 The coupling process

The incompressible fluid, solid tissue and acoustic solvers are explicitly coupled through the interface of the immerse boundary and the surface mesh of the tissue (Figure 2.2). The surface meshes of the solid tissue and the immerse boundary (vocal tract) at the interface do not need to be conformal. Interpolation is performed to provide more flexibility during mesh generation. In each time step, the displacement and velocity boundary condition of the fluid domain has been updated from the deformation of the solid tissue in the previous step (or set as zero during initialization). Then the incompressible flow is marched by one time step. The acoustic solver is then marched with the updated incompressible flow field as well as the displacement and velocity boundary condition. The incompressible flow pressure and acoustic perturbation pressure together would be used to update the nodal force applied on the vocal fold tissue. In the end of the time step, the solid solver is marched by one step with the updated traction boundary condition.

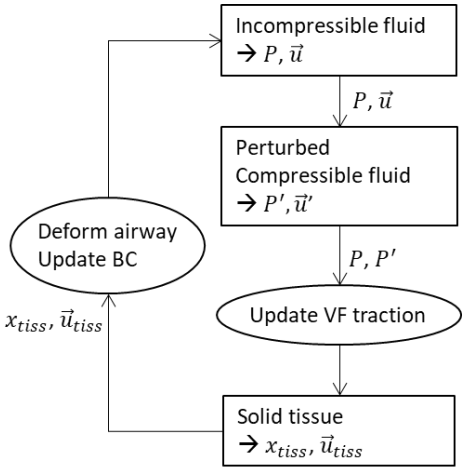


Figure 2.2 The coupling scheme of the solvers.

CHAPTER 3 VALIDATION AGAINST ANIMAL EXPERIMENTS

Due to the similarities in anatomical structures and vibration pattern, animal models have long been applied in the phonation studies. This chapter provides the validation of current numerical approach against the experiment measurements on a canine larynx and pigeon syringes. Model geometries are regenerated from the CT/MRI scan. Material properties have been inversely determined through an optimization method by minimizing the numerical and experimental difference of the static deformation. Numerical simulations follow the same setup with the experimental measurements and the key features and quantities have been compared.

The work about the validation on the pigeon syringes has been published as "Jiang, Weili, Jeppe H. Rasmussen, Qian Xue, Ming Ding, Xudong Zheng, and Coen PH Elemans. 'High-fidelity continuum modeling predicts avian voiced sound production.' *Proceedings of the National Academy of Sciences* (2020)."

3.1 The validation against a canine larynx

3.1.1 Experiment measurement

The experiment was performed by Dr.Khosla's group from Department of Otolaryngology-Head and Neck Surgery in University of Cincinnati Medical Center. One larynx was excised from a shared research Mongrel canine (weight 20kgs). The cartilages and soft tissues above the true vocal folds were removed to provide optical access to the glottis, and the mucosa was stitched to the edge of the thyroid cartilage, which was truncated at the folds' height. The trachea was cut 5 to 6 tracheal rings below the cricoid cartilage (~5cm). The larynx internal geometry at rest was obtained using magnetic resonance imaging scans (MRI scans).

One larynx was excised from a shared research Mongrel canine (weight 20kgs). The cartilages and soft tissues above the true vocal folds were removed to provide optical access to the glottis. The larynx internal geometry at rest was obtained using magnetic resonance imaging scans (MRI scans). The excised larynx was placed on an aerodynamic nozzle connecting to an air supply. The nozzle supported the larynx at the tracheal level. The vocal folds were adducted with two three-pronged mechanical supports that were inserted into the arytenoid cartilages, similarly to the setup featured in previous research by (Oren, Khosla and Gutmark, 2014b). Velocity measurements were conducted in the mid-membranous coronal plane (halfway between the anterior commissure and vocal process), using particle image velocimetry (PIV), similar to the setup in (Farbos de Luzan *et al.*, 2020) . The PIV camera was fitted with a Nikon 105 mm F/2.8 macro lens, which yielded a spatial resolution of 94.76 pixels/ mm. The PIV camera was placed at an angle of 40° relative to the x-y plane. Processing of the PIV data was done using DAVIS 10.5 software (LaVision GmbH) with a multi-pass decreasing pixel size (96 to 32) with a 75% overlap, which yielded a spatial resolution of 15.90 vectors/mm.

Projecting the laser sheet from above the larynx enabled to obtain intraglottal geometry and velocity measurements while the glottis was divergent. An electroglottograph (EGG) signal was used to determine the glottal closure instant, which was considered as the phase reference of every PIV instantaneous flow field, and used to trigger the PIV system (camera and laser). PIV measurement was performed nine loops to cover the phase time from 15 to 312 degree. Phase 0 was defined based on the EGG signal, which happened after the beginning of the closing but before the folds were completely closed. Each loop lasted 30 consecutive cycles and six different phases taken at 51.42-degree interval were acquired. The first phase times were shifted by 5 degrees in the next loop. The subglottal pressure (P_{SG}) was measured inside the nozzle using a pressure transducer (Honeywell, FPG). The sampling rates

for the pressure transducer, EGG, and TTL signals, were 20 kHz using a National Instrument data acquisition system (NI, PXIe-6356). The timing of the PIV, high-speed images, and data acquisition was synchronized using a shared reference clock.

3.1.2 Numerical model generation and determination of material property

The geometries of the vocal folds and related laryngeal cartilages were reconstructed from the MRI scans using the commercial software Mimics 16.0 (The Materialise Group, Leuven, Belgium) (Figure 3.1). The MRI scan had a resolution of 0.2mm in all three directions. The body-cover labelling scheme was adopted to describe the vocal fold inner structure and each layer was manually segmented for each vocal fold. The laryngeal cartilages were reconstructed to identify the contact areas of the vocal fold tissues and cartilages, where the constraint boundary conditions were applied in the simulations. The medial surface of each vocal fold was 12.5mm in the longitudinal direction and 3.5mm thick in the vertical direction (Figure 3.1 (c)).

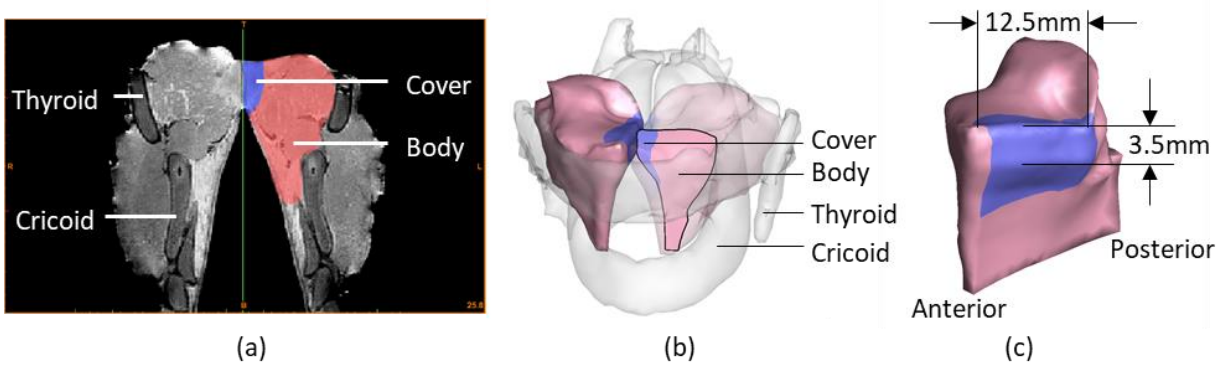


Figure 3.1 Model reconstruction from MRI scans. (a) A coronal plane of the larynx from the MRI scan, where the vocal fold tissue is annotated; (b) The reconstructed laryngeal model including the vocal fold tissues and related cartilages. The 3D geometry of the vocal fold on the right is made transparent to show the cover-body structure inside the vocal fold; (c) The dimensions of the medial surface of the vocal fold.

In this model, both the cover and body layers were modeled as transversely isotropic linear viscoelastic materials. Generally, five independent parameters are needed to describe this type of material, including the transverse Young's modulus (E_p), longitudinal Young's modulus (E_{pz}), transverse Poisson ratio (ν_p), longitudinal Poisson ratio (ν_{pz}) and longitudinal shear modulus (G_{pz}). In our model, the values of ν_p and ν_{pz} were set to be 0.9 and 0.0 (Alipour, Berry and Titze, 2000; Cook, Nauman and Mongeau, 2008), respectively, for both the cover and body, to model the material incompressibility. By considering that vocal fold vibration is mostly in the transverse plane due to longitudinal tension, the values of E_{pz} and G_{pz} were assumed to be 20 and 5 times greater than the magnitude of E_p, respectively, for both the cover and body layers. These treatments have significantly reduced the independent parameters of each layer, which now only includes E_p.

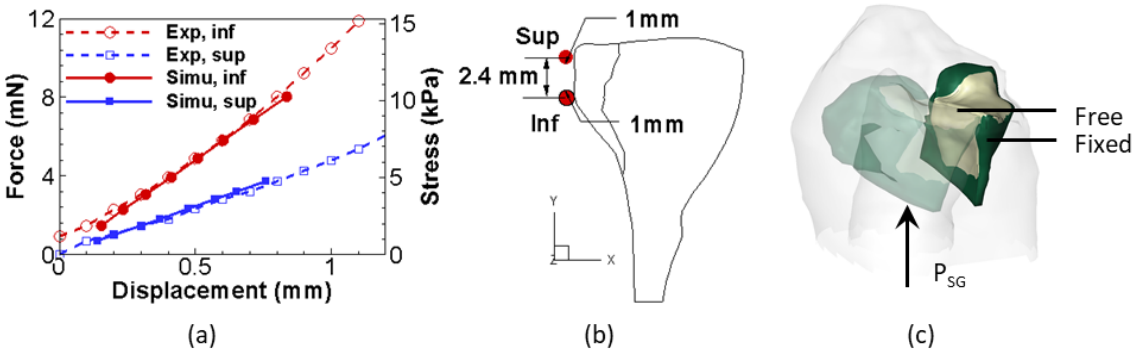


Figure 3.2 Material property and boundary condition. (a) Comparison of the force-displacement relationship obtained from the numerical and experimental (Oren *et al.*, 2014) indentations. (b) Schematic of the numerical indentation which was performed at the superior and inferior edges of the mid-coronal plane of the reconstructed vocal fold model. The indentation diameter was 1 mm. (c) Simulation setup and the boundary conditions. Surfaces denoted yellow were subjected to the aerodynamic loading and free to move. Surfaces denoted green were in contact with the cartilages and completely fixed during FSI simulations.

Past studies reported that stiffness of the cover layer increase from the superior to inferior aspect in both the human(Chhetri, Zhang and Neubauer, 2011) and canine (Chhetri and Rafizadeh, 2014; Oren *et al.*, 2014) vocal folds. This vertical stiffness variation can affect vocal fold dynamics in terms of increasing the opening and closing speed and promoting the vertical phase difference and divergent angle (Geng, Xue and Zheng, 2016). To include this feature in our model, another independent parameter---vertical stiffness gradient (VSG, kPa/mm), defined as the transvers Young's modulus gradient along the inferior-superior direction, was used to in the cover layer. In this way, the total independent parameters in our model include the transverse Young's modulus of the body layer ($E_{p,b}$, kPa), the transvers Young's modulus at the superior aspect of the cover layer ($E_{p,c}$, kPa) and VSG of the cover layer.

An inverse method has been developed to determine the three parameters from in-situ micro-indentation measurement. (Oren *et al.*, 2014) measured the force (pressure)-displacement relationships at both the superior and inferior edges of the mid-coronal plane of excised canine vocal folds. The data was provided for the present study and is shown in the dash lines in Figure 3.2(a). Note that the data was an average of 11 canine larynges, and the loading and unloading curves have been averaged to exclude the hysteresis effect. Following our previous work (Geng, Xue and Zheng, 2017), finite element method based numerical indentations were performed on the reconstructed model at the same locations as those in the experiment (Figure 3.2 (b)). A generic algorithm-based optimization process was developed to search for the optimal material parameters to match the numerical force-displacement curves with the experimental ones. The obtained numerical force-displacement curves were shown in the solid lines in Figure 3.2 (b). Note that the experimental curves reflected a nonlinear material behavior which cannot be represented by the linear material model in our method. However,

the nonlinearity only became evident when the strain was larger than 0.2 (0.8mm deformation). This corresponds to a stress larger than 4.7kPa at the superior edge and 10.2 kPa at the inferior edge. These values far exceed the possible pressure loadings on the vocal folds during vibrations. Thus, the numerical indentation was conducted up to the strain of 0.2. The values of the three material parameters obtained from the optimization were E_b= 4.76kPa, E_c= 1.33kPa and VSG= 0.43kPa/mm. These values were at the same order of magnitude as the experiment measurements reported in (Chhetri and Rafizadeh, 2014). Table 3-1 summarizes the material properties of each layer of the vocal fold in our model.

Table 3-1 Material property of the vocal fold layers. For the cover layer, due to the existence of the VSG, values are listed only for the superior side of the vocal fold. VSG is the vertical stiffness gradient.

	E_p (kPa)	E_{pz} (kPa)	G_{pz} (kPa)	ν_p	ν_{pz}	VSG (kPa/mm)
Body	4.76	95.24	23.81	0.9	0.0	-
Cover	1.33	26.70	6.68	0.9	0.0	0.43

In the simulations, the airflow was supplied by a subglottal pressure through the trachea of the larynx and exhausted to the open environment without a supraglottal tract (same as the experiments). Two subglottal pressures of 1.2 kPa and 1.8 kPa, corresponding to the low and high pressure conditions, were tested in the experiments and simulations. In the rest of the text and figures, SIM_L and SIM_H denote the high-pressure and low-pressure simulations cases, respectively, and EXP_L and EXP_H denote the high-pressure and low-pressure experimental cases, respectively. The kinematic viscosity of the air flow was $1.65\times10^{-5}m^2/s$ and the density was $1.15kg/m^3$. A grid-independence test of the Cartesian mesh of the flow solver was conducted on a coarse mesh with the smallest grid size of

0.08mm and a fine mesh with that of 0.04mm. The two meshes showed the maximum difference of about 5% at the maximum flow rate. Considering that the fine mesh increased the CPU time remarkably by about 8 times, the coarse mesh was used for the rest of the simulations. This resulted in a 128×128×128 non-uniform Cartesian grid setup.

In the vocal fold model, the superior, medial, and inferior surfaces (yellow surface in Figure 3.2 (c)) were the places subjected to the aerodynamic loading and free to move. The other surfaces (green surface in Figure 3.2 (c)) were connected to the cartilages and so fixed. The density of the tissues was 1.04×10^{3}kg/m^{3}. Each vocal fold model was discretized using 20643 tetrahedron elements. A hard-wall contact model was employed to mimic the vocal fold collision Moreover, a small artificial gap (0.16mm) between the vocal folds was enforced even during full glottal closure. It was for the success of the flow solver. The artificial gap resulted in small amount of leakages in the simulations. A small-time step of $5.5\text{x}10^{-4}$ ms was applied on both the flow solver and the solid solver, which was determined by the CFL stability constraint. The simulations were performed using 128 processors on the XSEDE supercomputer (Towns *et al.*, 2014). The computational time was about 5200 CPU hours / cycle for the low-pressure case and 6200 CPU hours /cycle for the high-pressure case.

3.1.3 Results and discussion

3.1.3.1 Glottal exit flow rate

Each simulation was carried on until about 6 steady cycles were obtained. The PIV measurement provided the detailed velocity field at the mid-coronal plane of the glottis, from which the planar flow rate at the glottal exit can be calculated. We computed that same flow rate at the same plane and same glottal exit position in the numerical simulations. Figure 3.3 compares the waveform of the flow rate between the experiments and simulations. The waveform is plotted in a dimensionless space in which

the flow rate is scaled by the peak flow rate and the time is scaled by the total duration of open glottis. Figure 3.3 shows a general agreement between the experiments and simulations that the flow rate rose very slowly at the beginning and then rose quicker and steadily in the rest of the opening phase until it was near the peak. During glottal closing, the flow rate dropped much quicker than its rising and maintained a steady rate during most of the closing time.

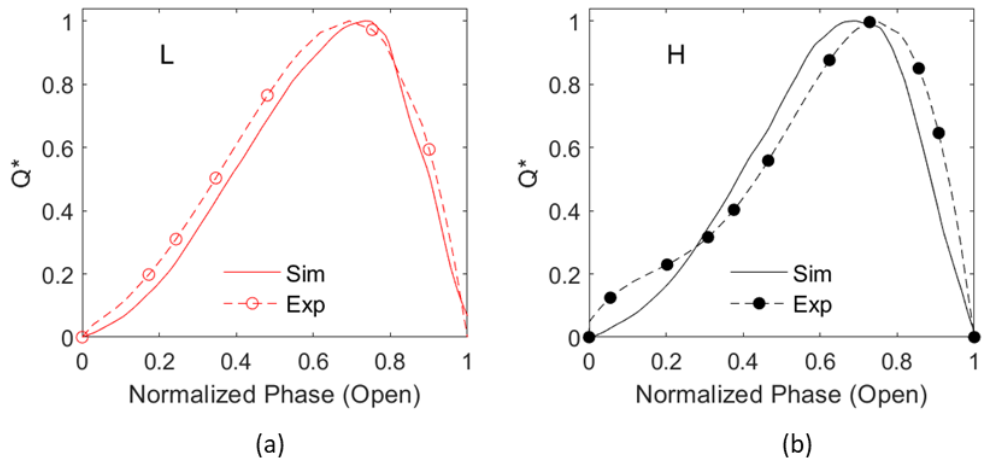


Figure 3.3 Comparison of the phase-averaged planar flowrate waveform. The flow rate was scaled by the peak flow rate and the time was scaled by the total duration of open glottis in each case. (a) is the low subglottal pressure condition and (b) is the high subglottal pressure condition.

Figure 3.4 compares several flowrate waveform related quantities between the experiments and simulations, including the open quotient, skewness quotient, frequency, maximum flow rate and mean flow rate. The values of the low- and high- pressure cases are denoted by the circle and square scatters, respectively. The bar plot denotes the mean value of the two pressure conditions. The open quotient was the ratio of the total duration of open glottis to the period, and the skewness quotient was the ratio of the duration of flow rate increase to the duration of decrease. On the mean value level, the simulations had an accurate prediction on the open quotient and skewness quotient. But the

simulations under-predicted the maximum and mean flow rates and over-predicted the frequency, suggesting that the elasticity of the vocal fold tissues may have been over-estimated. Note that the material parameters of the vocal fold tissues were determined based on the averaged indentation data from 11 canine larynges (Oren *et al.*, 2014). The real elasticity of the vocal fold tissues from which the model was reconstructed was not available. This is likely to cause the discrepancy. In the experiment measurements, the fundamental frequency increased with the subglottal pressure. However, it barely changed with the subglottal pressure in the simulation. This difference could be resulted from the linear assumption of the material property in the simulation.

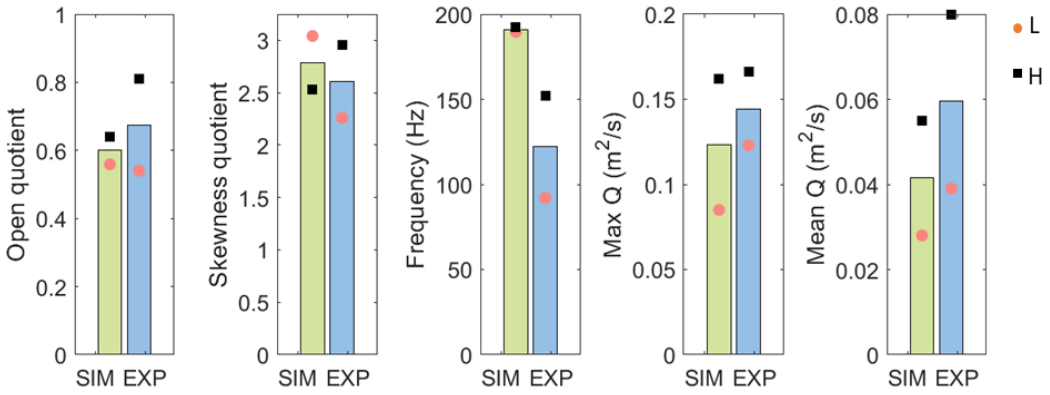


Figure 3.4 Comparison of flowrate waveform related parameters. The open quotient, skewness quotient, frequency, maximum flow rate and mean flow rate are compared. The values of the low- and high-pressure cases are denoted by the circle and square scatters, respectively. The bar plot denotes the mean value of the two pressure conditions.

3.1.3.2 Vocal fold vibration

Figure 3.5 shows the mid-coronal profile and top view of the vocal folds at two different time instants during one cycle in the high-pressure simulation case. The glottis was opening and closing at the two time instants, respectively. The changes of the mid-coronal profile clearly indicated a mucosal wave

traveling from the inferior to the superior, which resulted in a convergent glottal shape during opening and divergent glottal shape during closing. From the top view, the vocal folds vibrated in-phase in the anterior-posterior direction with the maximum amplitude occurring at the mid-coronal plane. The low-pressure cases presented the similar vibration pattern.

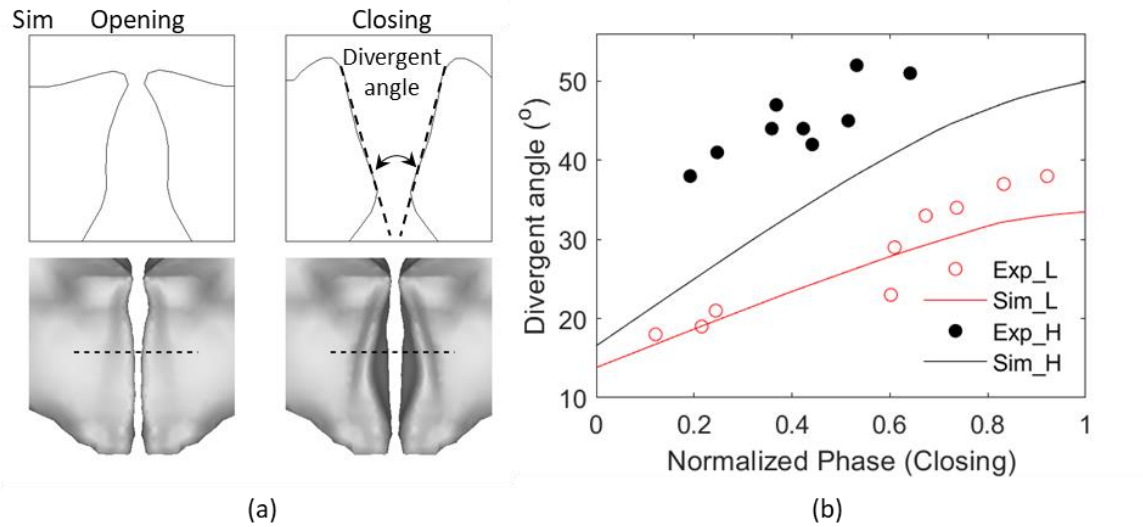


Figure 3.5 Comparison of the glottal angle. (a)The mid-coronal profile and top view of the vocal folds during vocal fold opening and closing in one cycle of the high-pressure simulation case. (b) Comparison of the time history of the glottal divergent angle during glottal closing between the experiments and simulations. The time was scaled by the total duration of glottal closing in each case.

The PIV measurement captured the vocal fold medial surface when the glottis presented a divergent shape, which has allowed an estimation of the glottal divergent angle. We also computed the glottal divergent angles during glottal closing in the numerical simulations. Figure 3.5 (b) plots the time history of the glottal divergent angle during glottal closing in the experiments and simulations. The time was scaled by the total duration of glottal closing. Note that the PIV measurements are phase-locked measurements and the glottal angle showed some level of dispersion due to cycle-to-cycle kinematic

variations. In the low-pressure cases, the simulation predicted the angles close to those in the experiments. Between 0-0.6 closing phase, the angles in the experiment and simulation were quite close. After 0.6 closing phase, the angles in the simulation were about 5^o smaller than those in the experiment. In the high pressure case, the divergent angles in the simulation were consistently about 10^o smaller than those in the experiment. The fact that the simulations under-predicted the divergent angles may be because of the under-predicted vocal fold vibration amplitudes due to the over-predicted elasticity as mentioned above. Studies reported that larger vibrations caused larger glottal angles (Oren, Khosla and Gutmark, 2019). Despite of these differences, similar dynamics of the glottal angle were observed between the simulations and experiments. First, the divergent angle continuously increased during glottal closing with the maximum angle occurring at the end of glottal closing. This suggested that the inferior had a higher closing speed than the superior edge through the entire closing phase. Second, the changing rate of the glottal angle with the dimensionless time was nearly linear in both the simulations and experiments. Third, a higher subglottal pressure generated higher divergent glottal angles.

3.1.3.2 Intraglottal flow fields

The PIV measurement provided detailed planar intraglottal velocity field, which allowed a thorough validation of the intraglottal flow simulations. Due to the accessibility, the PIV measurement is only possible when the glottis is in a divergent shape. Therefore, the measurement was only during glottal closing and was taken at the mid-coronal plane. We found that the low-pressure and high-pressure cases generally showed similar flow fields and dynamics. Therefore, the comparison between the experiments and simulations is only demonstrated on the high-pressure case here.

The phase-averaged flow field was calculated by averaging the flow field at the same phase over the steady cycles. It eliminated the cycle-to-cycle variation. Figure 3.6 (a)&(b) show the contours of the phase-averaged y-velocity (vertical velocity component) at two phases at the mid-coronal plane of the glottis in the high-pressure experiment (Figure 3.6 (a)) and simulation (Figure 3.6 (b). The two phases were at 0.45Tc and 0.74Tc, respectively, representing middle and late closing phases. Tc was the phase time normalized by the duration of glottal closing. In both the experiment and simulation, it was observed that, in both phases, the flow separated from the vocal folds walls at a location downstream of the minimum glottal width; after the separation, a jet was formed, and then the high-velocity core of the jet entrained the ambient flow, resulting in a recirculation zone between the jet and vocal fold wall at each side.

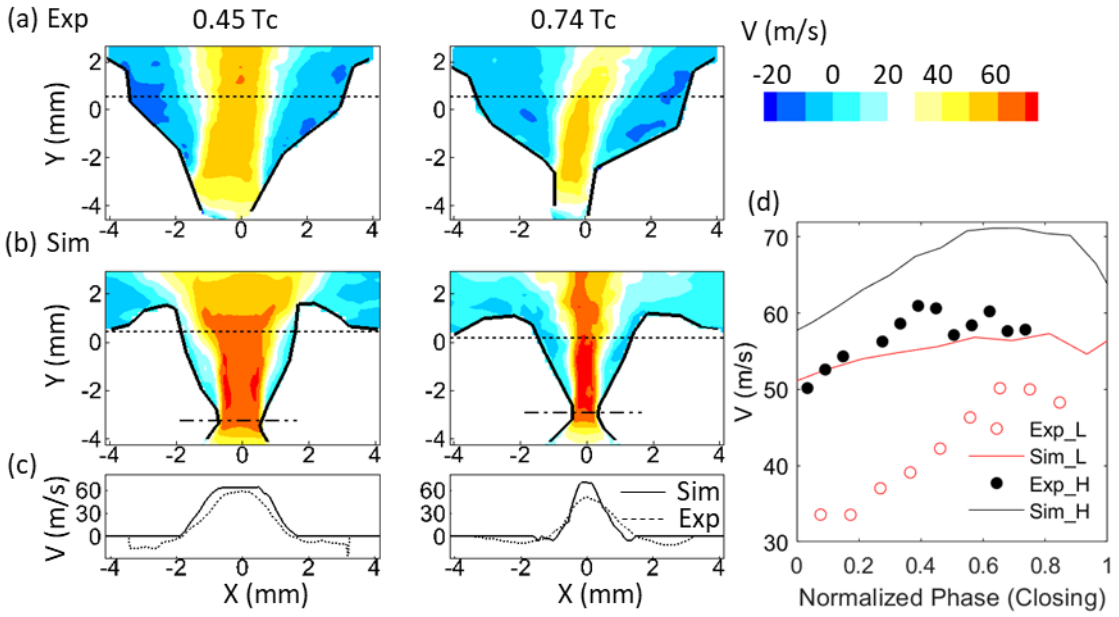


Figure 3.6 Comparison of the intraglottal flow velocity. Contours of the phase-averaged y-velocity (vertical velocity) at two phases at the mid-coronal plane of the glottis in the (a) experiment and (b) simulation with the high subglottal pressure. The two phases were at 0.45Tc and 0.74Tc, respectively,

representing middle and late closing phases. Tc was the phase time normalized by the duration of glottal closing. The location of the minimum glottal width in the simulation is denoted using the dash-dot line. (c) Profiles of the vertical velocity across the glottal width at the location of $\tilde{Y}$=0.8 corresponding to the velocity fields in (a) & (b). $\tilde{Y}$=0 and $\tilde{Y}$=1 correspond to the inferior and superior edge of the glottis, respectively (d) Time history of the maximum glottal jet speed at the mid-coronal plane of the glottis. The phase time was scaled by the duration of glottal closing.

The corresponding profiles of the vertical velocity across the glottal width at the location of $\tilde{Y}$=0.8 (dashed line in Figure 3.6 (a)(b)) were plotted in Figure 3.6 (c). $\tilde{Y}$=0 and $\tilde{Y}$=1 correspond to the inferior and superior edge of the glottis, respectively. In both the simulation and experiment, the velocity profiles showed a bell shape with negative velocities near the vocal folds walls, demonstrating the jet core and flow entrainment, and the jet width continuously decreased during glottal closing. If we define the jet width as the distance between the two locations where the velocity was 50% of the maximum velocity, the jet width decreased from 1.9mm to 1.5mm in the experiment and from 2.4mm to 1.0mm in the simulation from 0.45Tc to 0.74Tc.

Figure 3.6 (d) shows the time history of the maximum glottal jet velocity at the mid-coronal plane of the glottis in the experiment and simulation. The time was scaled by the duration of glottal closing in each case. The simulation generally predicted a higher jet velocity than the experiment, which is especial evident in the low-pressure case. It was likely due to that the simulation under-predicted the vocal fold vibration amplitude and so the glottal opening. Both the simulation and experiment revealed that the maximum jet speed continued to increase at early time of glottal closing despite that the flow rate already started to decrease. The jet velocity was highest in the middle or late of the closing phase. A close examination found that the delay of the maximum glottal jet velocity comparing to the flow rate

was due to the fact that the decrease of the glottal area, which caused the velocity to increase, was much faster than the decrease of the flow rate at early glottal closing.

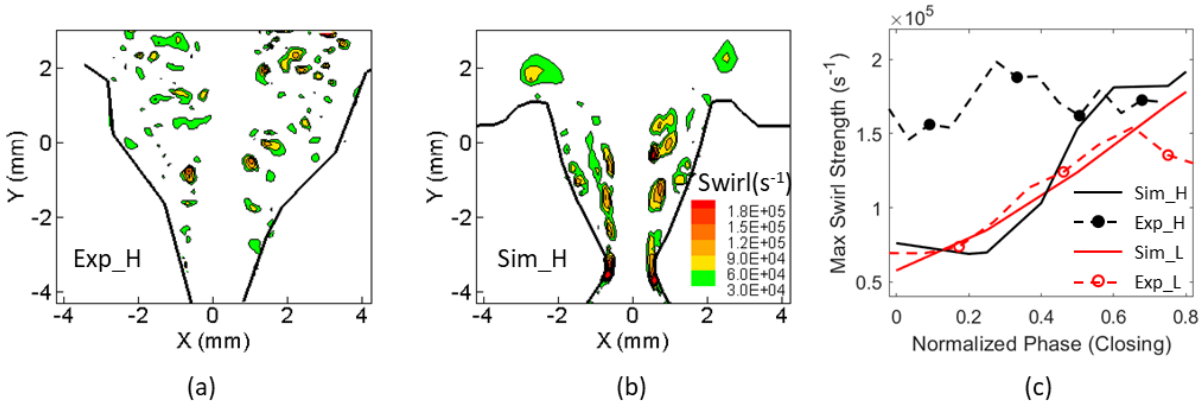


Figure 3.7 Comparison of the intraglottal swirl strength. Time-instantaneous intraglottal swirl strength at 0.45Tc, where Tc is the duration of glottal closing, in the (a) experiment and (b) simulation with the high subglottal pressure. (c) Variation of the maximum swirl strength during glottal closing. The phase time was scaled by the duration of glottal closing of each case.

Figure 3.7 (a)&(b) show the contours of the swirl strength of the vortical structures in a time-instantaneous flow field at 0.45Tc in the experiment and simulation with the high subglottal pressure. The swirl strength was calculated as the image part of the eigen value of the local velocity gradient tensor (Adrian, Christensen and Liu, 2000). In both the experiment and simulation, a large number of small vortical structures were generated between the jet and vocal folds walls, and it's likely due to the shear layer instability due to the jet separation. The swirl strength of the vortical structures was at the same order of magnitude between the experiment and simulation. Figure 3.7 (c) compares the time history of the phase averaged maximum swirl strength during glottal closing between the experiments and simulations at both subglottal pressures. The time was scaled by the total duration of glottal closing. Note that the experimental data toward the very late closing was not available due to experimental

difficulties. It was observed that the predicted value and dynamic variation of the swirl strength in the low-pressure case was very close to those in the experiment. The swirl strength continued to increase during most of the closing phase. Such pattern was also observed in the high-pressure simulation. The high-pressure experiment, however, showed a fluctuation pattern of the swirl strength around a constant during the entire closing phase. But we also noted that, in a previous PIV measurement on excised canine larynges(Oren, Khosla and Gutmark, 2014a), the swirl strength increased with the closing phase and the subglottal pressure.

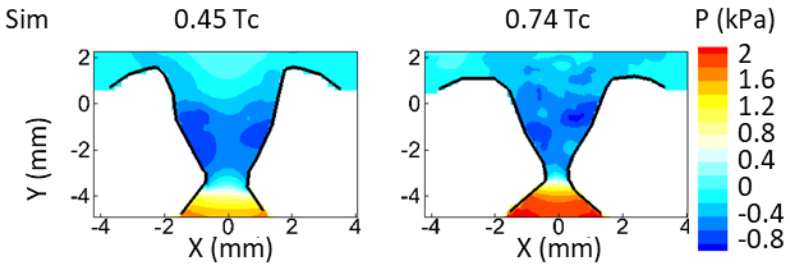


Figure 3.8 The pressure contours of the phase-averaged flow fields at the mid-coronal plane in the high-pressure simulation case.

Figure 3.8 shows the pressure contours of the phase-averaged flow fields at the mid-coronal plane in the high-pressure simulation case. It was observed that the pressures were mostly negative inside the glottis with the value at the same order of magnitude as the subglottal pressure. It is generally thought that the air inertia in the supraglottal tract is the primary mechanism generating negative pressures inside the glottis during glottal closing, and if there is no supraglottal tract, the intraglottal pressure after flow separation would be around the ambient pressure based on the Bernoulli's principle. In our simulations, there was no supraglottal tract, but significant negative pressures were observed inside the glottis throughout the closing phase. This indicated that these negative pressures were

generated by the flow viscous effect. In fact, it was noticed that the negative pressures were more significant in the recirculation zones where the intraglottal vortices were generated (Figure 3.7). It was also noticed that the negative pressures were most significant at the late closing phase (Figure 3.8), which was consistent with the growth of the swirl strength (Figure 3.7 (c)).

3.2 The validation against pigeon syringes

3.2.1 Experiment measurement

The experiment was performed by Prof.Elemans' group from Institute of Biology in University of Southern Denmark. The subjects were six adult domestic pigeons (*Columba livia*) obtained from local breeders. All procedures were carried out in accordance with the Danish Animal Experiments Inspectorate (Copenhagen, Denmark). Subjects were euthanized by overdosing with Isoflurane (Baxter Medicals, Il, USA) and the syrinx was extracted from the primary bronchi up to including ~10 cm of trachea (Figure 3.9 (a)). The syrinx was mounted to tracheal and bronchial connectors in the experimental setup described in detail in (Elemans *et al.*, 2015; Rasmussen, Herbst and Elemans, 2018). In brief, this setup allows fine control of bronchial (p_b) and air sac (p_{as}) pressure, flow, temperature, and simultaneous high-speed imaging from sagittal and endoscopic tracheal views. Sound is recorded through a microphone placed 5-10 cm from the tracheal connector outlet on a 45° angle to avoid air jets from the tracheal outlet. In natural condition, the vibratory tissue in syrinx includes lateral vibratory mass (LVM) and medial tympaniform membrane (MTM). To decrease the control parameters, MTM was glued not to move leaving LVM the only vibratory tissue in the current setup. Tension of the LVM is adjusted by the pressure difference of p_b and p_{as}.

First, static deformation of the LVM was performed and measured to obtain the LVM tissue properties later (Figure 3.9 (b)). The LVM is assumed to be incompressible isotropic material based on

earlier histological sections (Elemans, 2004). Hence there is only one unknown material property: the elastic modulus (E). After mounting, the stress-stain relationship of the LVM tissue was measured using a static loading test. The transmural pressure (p_t), defined as $p_t=p_b-p_{as}$, was stepwise increased every 1.5s from 0-2 kPa (in increments of 0.1 kPa) corresponding to the in vivo pressure range (Elemans, Zaccarelli and Herzel, 2008). To prevent airflow through the syrinx, the trachea was closed. The static deformation of the LVMs has been imaged using a 1.2 mm flexible endoscope (Scholly, Germany) attached to a videokymographic (VKG) system (Videokymographic camera 2156, Cymo B.V., The Netherlands), which combines a high-speed linescan camera (7,200 lines/sec) with a full frame CMOS camera (25 frames/sec). The analog video output from the VKG was digitized together with the microphone and a synchronization signal using a video capturing device (Intensity Extreme, Black Magic Design, Australia). The digital kymograms were constructed perpendicular across the LVM. The edge position was quantified using outwards-in threshold detection of the edge. The length of dorsoventral glottis at mid LVM position was determined from each CT scan and allowed us to calibrate LVM displacement from the DKG in mm. For statistical purpose, five technical replicates were made for each transmural pressure setting.

Second, the time-resolved motion of the LVMs, flow, pressure and sound during phonation were quantified. The boundary condition is: p_b = 1.0 kPa, p_{as} = 0.5 kPa, where LVM oscillations reliably occurred (Elemans *et al.*, 2015; Rasmussen, Herbst and Elemans, 2018). The resulting positive transmural pressure mimics activation of the TL muscle (Elemans, Zaccarelli and Herzel, 2008). A powerful, stable light source (1700 Lumen white LED powered by PS23023, HQ Power amplifier, Belgium) trans-illuminated the syrinx from dorsal allowed us to capture LVM motion with a high-speed camera (MotionPro-X4, 12 bit CMOS sensor, Integrated Design Tools, Inc.; 4,000 frames/sec) mounted

on a stereomicroscope (M165-FC, Leica Microsystems). Before and after experiments, the mounted preparation was made a photo with a Leica DFC400 digital camera mounted on the M165-FC stereoscope. LVM shape was hand-traced in all specimens on 150-200 selected consecutive high-speed images (covering at least 5 full vibratory cycles) using a custom-made Matlab graphical user interface.

After experiments were completed, the geometries of the syringes were obtained through microCT scan, taking care to mimic the exact geometry of the syrinx during the experiments based on earlier photos. One preparation was omitted for further analysis as it twisted after fixation compared to the experimental situation.

A blinding procedure was employed that the modeling team (WJ, QX, XZ) was given access only to the DiceCT scan, static loading test (LVM displacement and transmural pressure) and asked to predict key performance traits of LVM kinematics (glottovibrogram) and acoustic waveforms during voiced sound production under listed posturing and pressure boundary conditions. After the simulations of all preparations were completed all experimental data was disclosed and analyzed.

3.2.2 Numerical model generation and determination of material property

To parameterize tissue elasticity for each subject, the finite-element LVM model was used to simulate LVM displacement as a function of p_t as obtained in the experiment for a range of elastic moduli values. Uniform bronchial pressure steps were applied on the inner LVM surface and the maximum LVM gap was obtained for each pressure step. For each preparation, different elastic modulus values generated by a genetic algorithm-based optimization method (Liu *et al.*, 2019) were randomly assigned to each static analysis solver, which resulted in a different displacement-transmural pressure slope. The elastic modulus of the LVMs for each preparation was determined in an iterative process by minimizing the difference between the slopes of LVM displacement versus p_t measured in the

experiment with the 3D mesh model (Figure 3.9 (c)). The elastic modulus values obtained through this process ranged from 1.8 to 4.0kPa (Table 3-2).

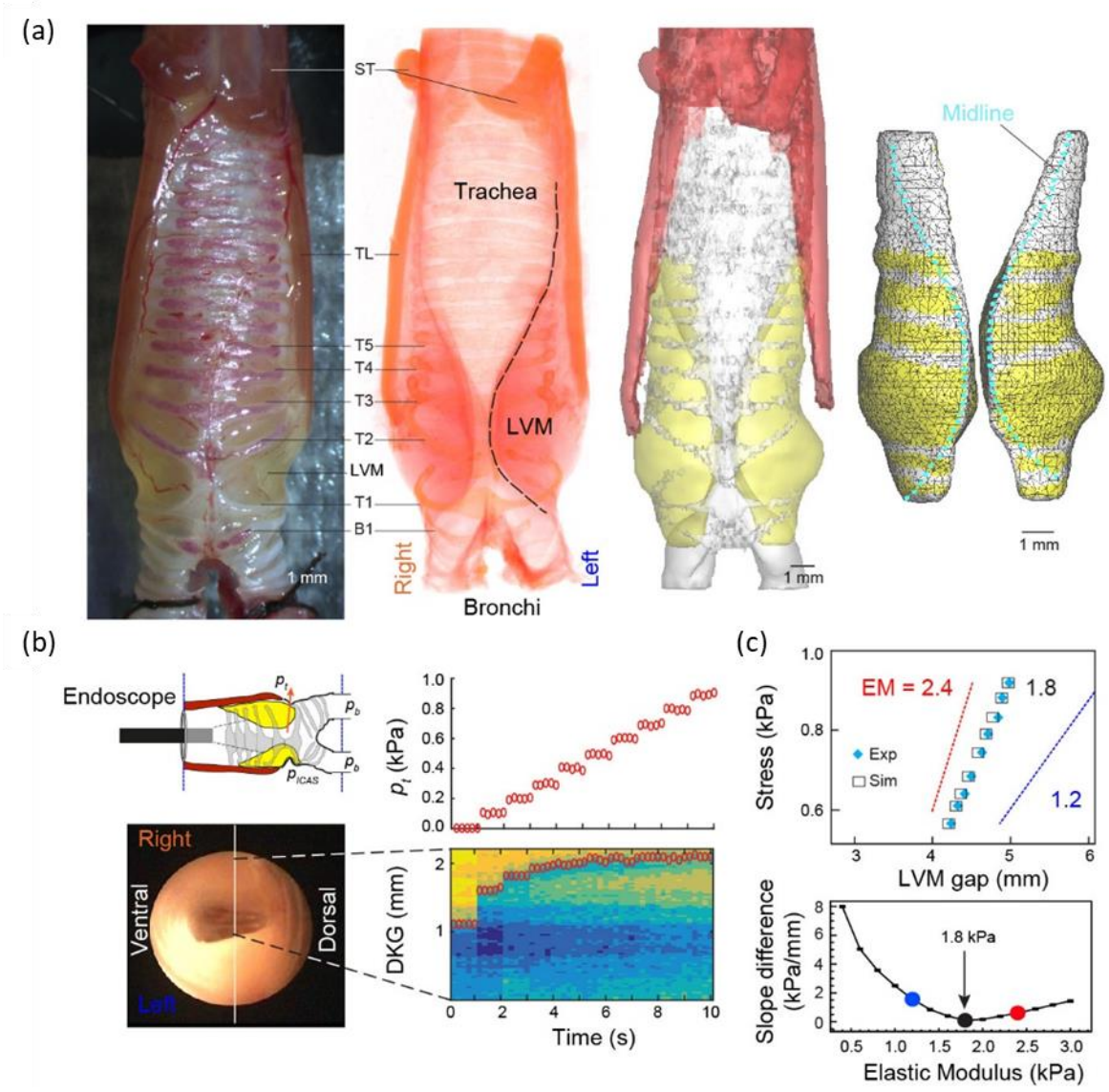


Figure 3.9 Parameterization of vocal organ geometry and tissue properties. (a) Workflow to parameterize FSAI model geometry (Subject P1), with from left to right: 1) photo of syrinx mounted in in vitro experimental setup, 2) voltex rendering of iodine contrasted microCT scan showing the bilateral Lateral Vibratory Masses (LVM), 3) 3D anatomy, and 4) Finite Element Mesh of LVM solid domain.

Yellow and white mesh elements on LVM outer surface have free and zero displacement boundary conditions, respectively. (b) Tissue properties were determined by a combined experimental and modeling approach. A stepwise increase of (top) transmural pressure (p_t) caused a stepwise sideways displacement of the LVMs (bottom) as viewed by digital kymogram (DKG) along the white vertical line in the endoscopic image (left). (c) The reverse engineering approach used to determine the LVM elastic modulus (EM) for each individual. Top: stress-gap width curve of experimental (blue diamonds) and modeled displacement for LVMs with EM=1.8 kPa (black squares). Also indicated are simulated aligned gap widths of EM=1.2 (blue line) and 2.4 kPa (red line). Bottom: Minimal difference between experimental and simulated data indicates occurs at EM = 1.8 kPa for this individual. B1, bronchial ring; T1,…, T5, tracheal rings; LVM, Lateral Vibratory Membrane; pt, transmural pressure TL, tracheolateralis muscle; ST, sternotrachealis muscle.

Table 3-2 LVM finite element model statistics and elastic moduli per individual.

	Nodes		Elements		EM
	Left LVM	Right LVM	Left LVM	Right LVM	kPa
P1	3988	3835	15737	15190	1.69
P2	1988	2278	8101	9838	1.8
P3	2686	2159	11402	8532	4
P4	4253	4551	17062	18436	2.24
P5	4620	4693	18657	19131	2.4

3.2.3 Simulation setup

The vocal organ mesh was immersed in a computational domain that was about 12×10×40 mm in dimensions and straight tubes were added to each bronchial inlet and tracheal outlet to avoid reverse flows. The domain was discretized with a non-uniform 64×48×128 Cartesian mesh with the highest grid density around the LVMs. A non-penetration non-slip boundary condition was applied at the airway wall. The part of the airway wall contacting the LVMs was the interface of the fluid-structure

interactions on which the deformation and velocity of the LVM surface were transferred to the fluid solver and the fluid pressure on the LVM surface was transferred to the solid solver. The air density was set to 1.1455 kg/m^3 at 37°C. We used a kinematic viscosity value of 6.6×10^{-5} m^2/s (at 37°C) - four times the normal value - to reduce high frequency turbulence in the simulations, which significantly reduces the computational costs (Zheng *et al.*, 2011). The Poisson ratio (v) was set to be 0.46 to avoid the singularity problem at v=0.5. The shear modulus (G) was obtained from the relationship of G=E/2/(1+v) for isotropic materials. Tissue density was assumed to be 900 kg/m^3, equaling to the density of fat.

The acoustic results were first calculated using the LPCE method as mentioned in 2.1.3. However, the calculation of the LPCE showed that the root-mean-value of the acoustic flow rate was only 0.9% of the incompressible flow rate, suggesting a weak coupling effect between the acoustic field and incompressible flow field. Therefore, the linear source-filter theory was applied which assumes that the sound radiation in the tract is linearly coupled with the source generation in the syrinx so that the acoustic pressure does not affect vibrations. Because the dominant sound source of voice is a monopole source, the far-field acoustic pressure was calculated as: $p' = \frac{\rho}{4\pi r}\frac{dQ}{dt}$, where p' is the acoustic pressure, ρ is the air density, r is the distance from sound source to the microphone and dQ/dt stands for the time derivative of flow rate.

The pressure boundary condition is p_b=1.0kPa and p_{as}=0.5kPa. A time step of about 1.0 μs was utilized in the fluid and solid solvers. Two artificial non-slip and non-penetrable collision planes were 80 μm off the medial plane to enforce a finite 160 μm minimum gap between the two LVMs during closure. The LVM position was not allowed to exceed the corresponding collision plane during collision, to prevent failure of the solver due to the non-conserved mass in each zone. For each individual, the

simulation was performed on 16 processors and took about 1.5 days per vibration cycle. At least five cycles were obtained per individual.

3.2.4 Results and discussion

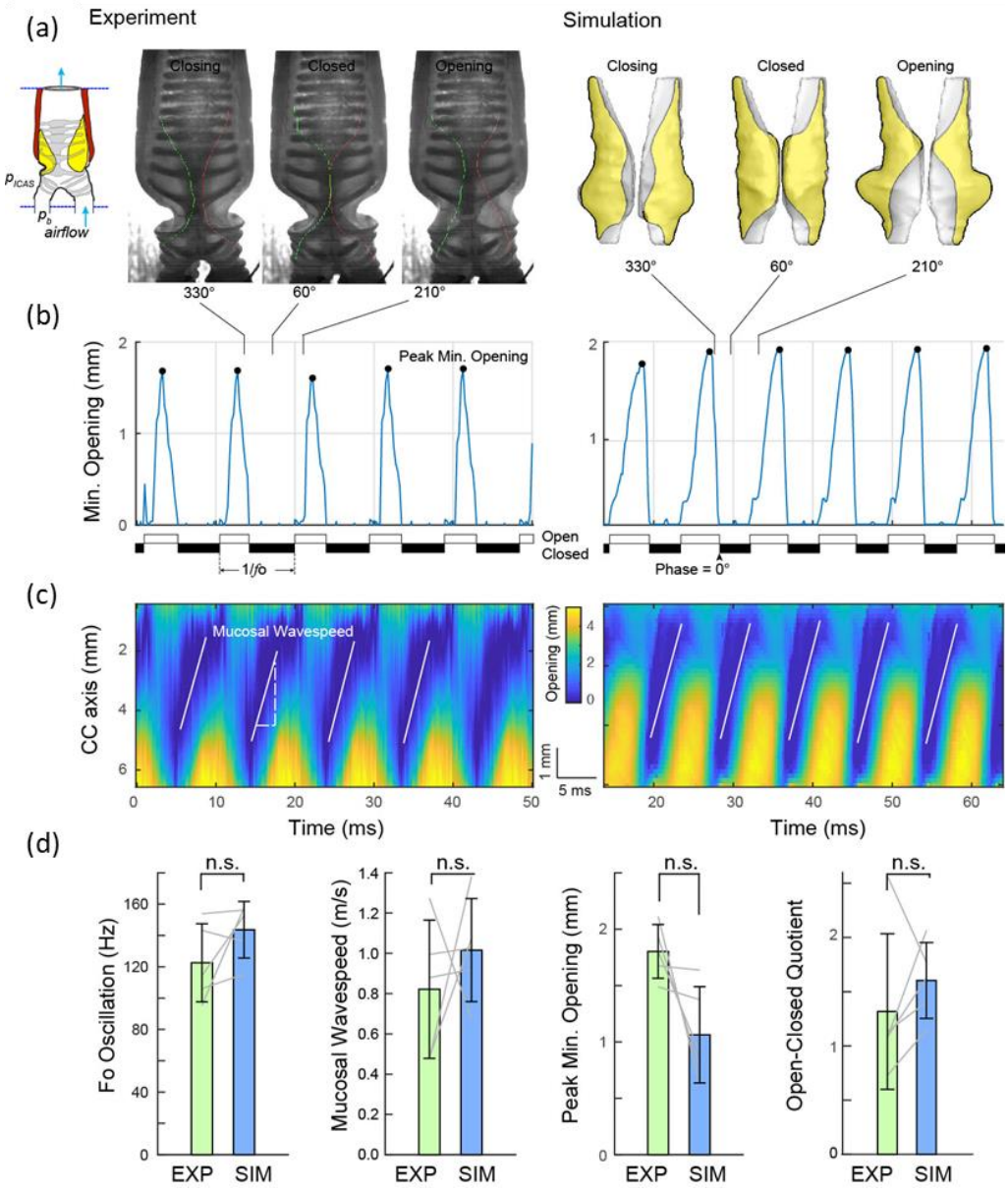


Figure 3.10 FSAI simulation accurately predicts key features of LVM kinematics. (a) Trans-illuminated syrinx in experiment (EXP, left) and simulation (SIM, right) showing the LVM shape in coronal plane at

330, 60 and 210° of the vibratory cycle, where 0° is defined as the first frame of full closure. (b) Glottal opening with white and black horizontal bars denoting open and closed phase. (c) Coronal glottovibrogram showing time-resolved glottal opening along caudocranial axis of 5 oscillations. The white lines indicate the mucosal wavespeed regression slopes based on the closed glottis. (d) Observed experimental data and simulation predictions are not significantly different for the above indicated four key kinematic parameters (See Table 3-3 for values and statistics).

Table 3-3 Comparison of experimental and simulation data.

	Simulation		Experiment		Heteroscedastic two-tailed paired t-test p =
	Mean ± S.D. n=5	Range	Mean ± S.D. n=5	Range	
F_0 (Hz)	143±18	115-158	123±25	95-154	0.18
Mucosal wave speed (m/s)	1.02±0.26	0.68-1.38	0.82±0.34	0.48-1.27	0.50
Peak minimal opening (mm)	1.06±0.43	0.62-1.64	1.80±0.24	1.49-2.11	0.06
Open closed quotient	1.6±0.3	1.1-2.1	1.3±0.7	0.7-2.5	0.11

Both the in vitro syrinx and corresponding individual FSAI models demonstrated self-sustained stable oscillations in all five cases (Figure 3.10). To quantifying the time-resolved vocal fold shape within oscillations, we took advantage of both the unique coronal view offered by the pigeon syrinx (Elemans *et al.*, 2015) and the lack of a dorsoventral vibrational component (Rasmussen, Herbst and Elemans, 2018), to quantify the time-resolved syringeal or glottal opening as a function of caudo-cranial position, i.e. a coronal glottovibrogram (Figure 3.10(c)).

Key vibratory and acoustic predictions by the simulations were compared to the behavior in the experiment. The glottovibrogram allowed for comparing four key parameters describing vibratory kinematics: fundamental frequency of the vibration (F_o), the speed of the mucosal wave, peak of the

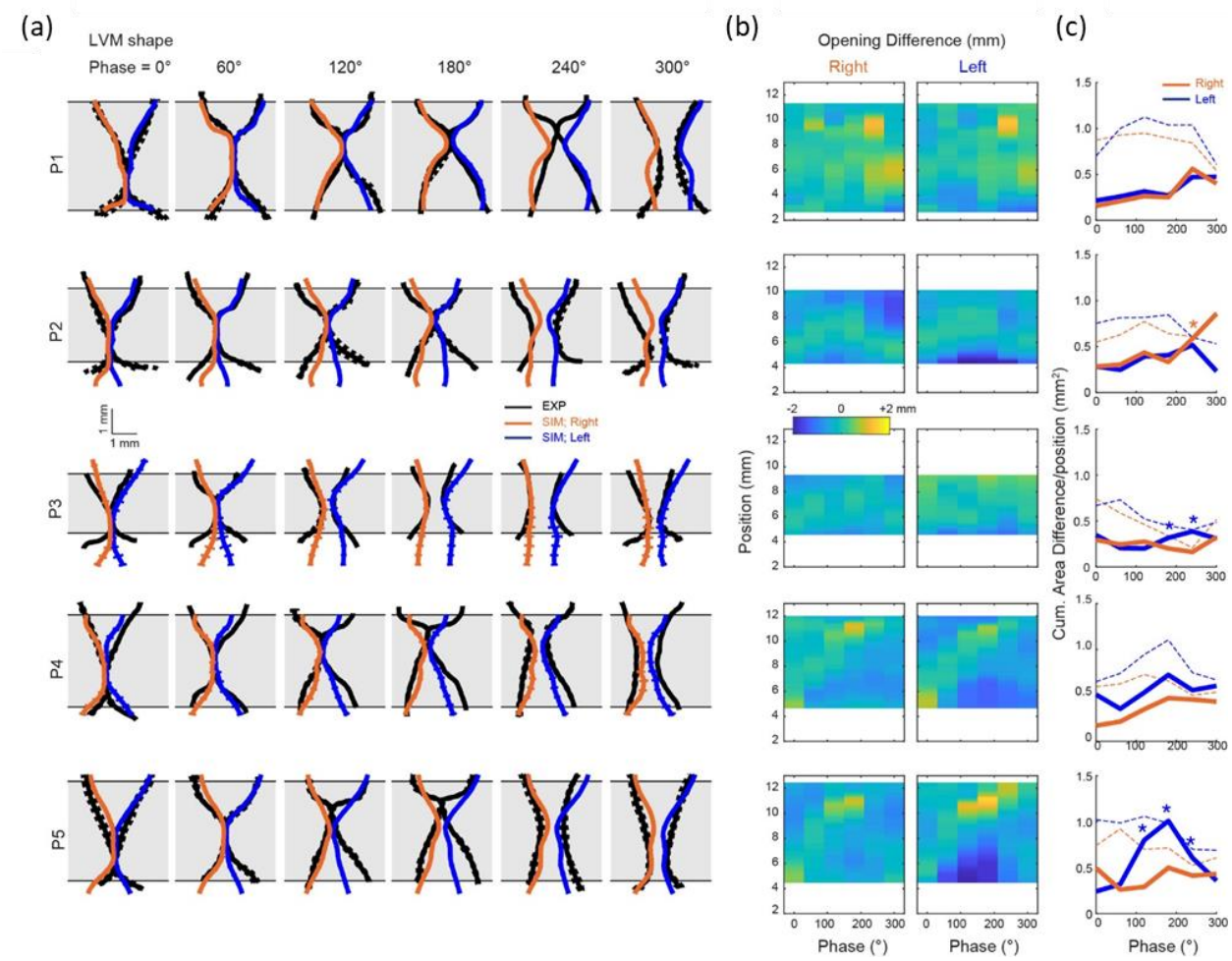


Figure 3.11 FSAI simulation accurately predicts within-cycle LVM shape. (a) Bilateral LVM shape of experiment (EXP) and simulation (SIM) in coronal plane at 0, 60,120, 180, 240, 300 degrees averaged over 5 cycles (mean (solid lines) ± S.D (dotted lines). (b) Lateral position difference between experiment and simulation left and right LVMs and (c) unsigned cumulative area difference between experiment and simulation shows that the predicted shapes are very similar to the observed shapes. The most pronounced discrepancies occur during late closed/early opening. Dotted lines are mean. Asterisks indicate significant difference ($p>0.01$) between simulation data and 1000x bootstrapped experimental LVM shapes for 2-sample Kolmogorov-Smirnov tests.

minimal glottal opening, and the open-closed quotient. The open-closed quotient is the ratio between duration of open and closed glottis. At group level, all predicted values were not significantly different from the experimental values (Table 3-3).

The LVM shapes between experiment and simulation at fixed phases within an oscillatory cycle were further compared (Figure 3.11). During the closed phase (0-120°) the simulated LVM shape matched the experiments very well and was never significantly different from the bootstrapped shape ($p<0.01$). The only observed significant ($p>0.05$, 2-sample Kolmogorov-Smirnov test) discrepancy occurred during late closed/early opening in three subjects on one side, where the LVM mass tended to move ~0.5 mm more cranial (superior in human anatomical terminology) in experiments compared to simulations (Figure 3.11 (b)(c)). Taken together, the model accurately predicted key kinematic parameters of LVM motion.

Lastly, two key acoustic parameters specifying a sound source – in addition to f_0 (Figure 3.10(d)) -, namely source level and spectral slope were compared and found to be with no significant difference (Figure 3.12(a)). Because the simulated LVM vibratory kinematics matched the experiment measurement, parameters that could not be quantified in the current experiments could be calculated through the numerical model, such as spatiotemporal pressure and flow velocity distributions over the vibratory cycle (Figure 3.12 (b)-(e)). The convergent LVM shape during opening causes high glottal pressures (~0.9 kPa) that transfer 23.4±15.7 μJ of positive energy (N=5) from flow into LVM (Figure 3.12(f)), facilitating opening. When maximum opening is reached at 300° phase, the LVMs are straight (0° angle in Figure 3.12(e)) reducing glottal pressure. During early closing the inferior LVM edge is moving inwards causing an energy transfer of 7.2±2.2 μJ back into the flow. Consecutively, the LVMs close the glottis by moving together in a divergent shape causing rapid pressure reduction driven by

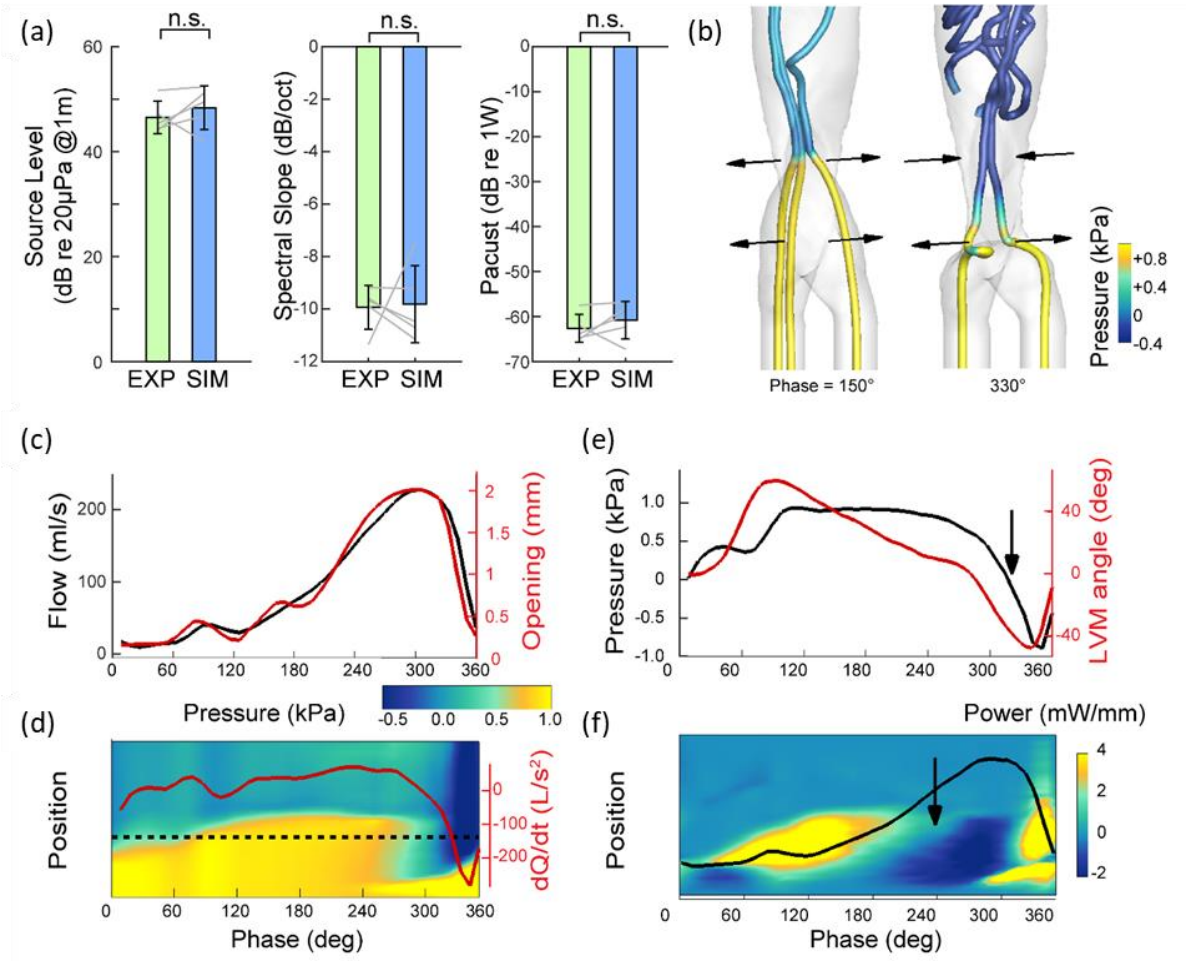


Figure 3.12 Numerical predictions of acoustics and spatiotemporal pressure and power profiles. (a) The simulations accurately predicted key acoustic parameters: source level, spectral slope and acoustic power. (b) 3D flow inside the syrinx as indicated by flow streamlines (pressure color-contoured). Horizontal arrows indicate the motion direction of the LVMs. (c) Flow rate Q (black line) correlates strongly with glottal opening (red). (d) Spatiotemporal pressure distribution along the airway centerline over a cycle with superimposed dQ/dt (red solid line) as a proxy for sound pressure. (e) Glottal pressure (black line) evaluated at the horizontal dotted line in (d) with LVM angle (red line). (f) Spatiotemporal power transfer distribution from flow to LVM evaluated along the airway centerline with superimposed flow Q (black solid line).

elastic forces (Arrow in Figure 3.12(e)). Furthermore, flow inertia in the trachea causes negative pressures (-0.52±0.23 kPa) near the glottis exit prior to full closure (blue region in Figure 3.12(d)), facilitating closing. Interestingly, another positive energy transfer to LVMs is observed nearly the end of the cycle. Thus it show that two primary factors contribute to the pressure asymmetry during vocal fold opening and closing that drive self-sustained oscillation: i) an alternating convergent/divergent medial surface profile and ii) airflow inertia, corroborating earlier model predictions (Titze, 1988) and measurements on (hemi)larynges (Alipour and Jaiswal, 2008) in mammalian voice production.

3.3 Summary

The validation of current numerical approach against the experiment measurement on a canine larynx and pigeon syringes has been provided. In both studies, the material properties were determined through an optimization-based inverse process using experimental indentation measurement in the canine case and the static deformation measurement in the pigeon case.

In the canine case, the simulations and the experiments results are compared with respect to the glottal flowrate waveform, glottal jet dynamics, vocal fold divergent angles, intraglottal velocity fields, vortical structures and swirl strength. The simulation result shows higher fundamental frequencies and lower glottal flow rate, which might be due to the overestimated tissue elasticity. Some features, such as the increased frequency with increased subglottal pressure was not observed in the simulation result due to the linear assumption of the material property. Regarding the intraglottal flow fields, the model predicted the key features observed in the experiments, including the similar glottal jet dynamics and intraglottal recirculation zones in the phase-averaged flows and similar intraglottal vortical structures and swirl strength in the instantaneous flows. The limitation of the validation of the canine case is that the material measurement was an average of 11 canine larynges, the obtained

material parameters may be different from the exact properties of the specific larynx model used in the present study.

The pigeon study provides a strict one-to-one comparison in that the geometry, material property and dynamic measurement all come from one single subject. The numerical model accurately predicted key performance traits of tissue motion and acoustics driven solely by physiological parameters (static geometry, tissue elasticity, and boundary conditions) and without optimization of either geometry or material properties parameterization on dynamic performance (Chang *et al.*, 2016). While the numerical model implementation itself is complex, the inputs are simple and have, most importantly, directly measurable physiological, material, and geometrical properties. Measurements of vocal fold material properties, initial configuration (Miri, 2014; Dion *et al.*, 2016), initial stress, and detailed flow-induced 3D vocal fold motion (Vahabzadeh-Hagh, Zhang and Chhetri, 2017b) have been achieved separately in human and mammalian model systems, but complete physiological data sets have not been obtained in these clades, nor in any birds, in single individuals nor consecutively used to thoroughly test numerical model predictions in a blinded approach. Recent studies encouragingly suggested that realistic continuum 3D models lead to more robust vocal fold dynamics compared to 1D and 2D models(Zheng *et al.*, 2011). The current data shows that realistic continuum 3D models also lead to accurate predictions. Moreover, it presents strong support for the current numerical approach as a critical first step toward the endeavor of integrating in vitro, ex vivo, and in vivo experimental data with brute force computational approaches into a causal model of motor control of voiced sound production applicable to birds and mammals.

CHAPTER 4 FLUID-STRUCTURE-ACOUSTIC INTERACTION SIMULATION IN SIMPLIFIED HUMAN LARYNX AND VOCAL TRACT SYSTEM

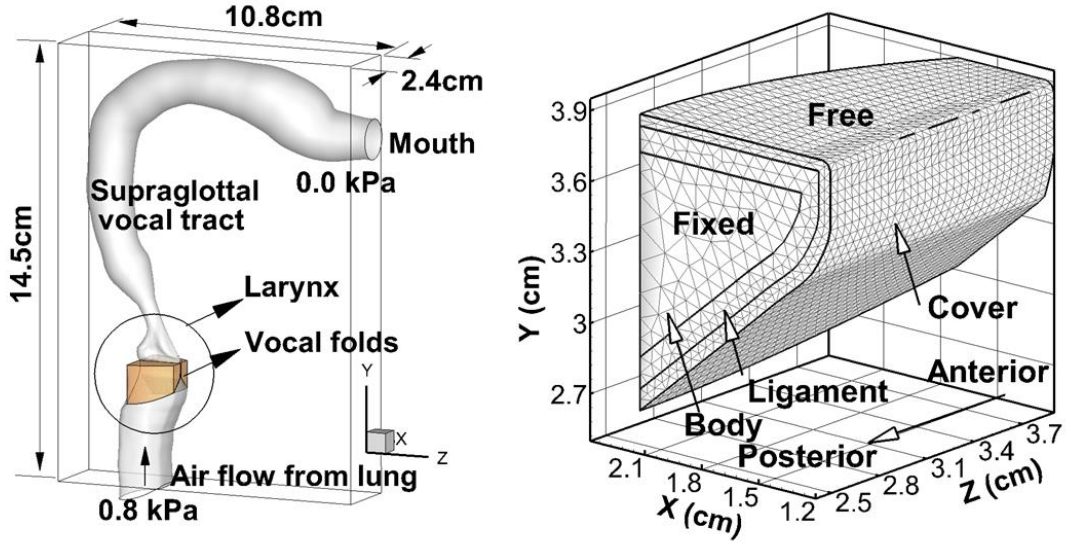
This chapter provides the application of the numerical approach on a simplified human vocal fold and vocal tract system, which will serve as the baseline case in the later chapters. The study in this chapter has been published in "Jiang, Weili, Xudong Zheng, and Qian Xue. 'Computational modeling of fluid–structure–acoustics interaction during voice production.' *Frontiers in bioengineering and biotechnology* 5 (2017): 7."

4.1 Model and simulation setup

The geometric model of the simulation is shown in Figure 4.1 (a). The geometry of the larynx was roughly reconstructed from a thin-slice CT scan of the larynx of a 30-year-old male subject (Zheng *et al.*, 2009). The geometry of the vocal folds was constructed based on the mathematical model proposed by (Titze and Talkin, 1979), which has considered the three-dimensional shape of the vocal fold including the anterior-posterior variation. The cross-section area of the supraglottal vocal tract was taken from an in-vivo based neutral vowel model proposed by (Story, 2005), and it was superimposed onto a realistic airway centerline from the in-vivo MRI measurement (Story, Titze and Hoffman, 1996) to generate the supraglottal tract model. The length of the supraglottal tract was 17.4cm and the length of the subglottal tract was 3.05cm. The vocal tract generally did not move except the place that contacted with the vocal folds.

The vocal fold was divided into three layers including the cover, ligament and muscle (Hirano, Kurita and Nakashima, 1981). Longitudinal variation in layer thickness exists (Hirano, Kurita and Nakashima, 1981) but has been shown to have a negligible effect on vocal fold vibrations (Xue *et al.*, 2011). Therefore, each layer was assumed to be longitudinally invariant in the current model. The

thickness of the cover and ligament layer was 0.5mm and 1.1mm, respectively, adopted from (Titze and Talkin, 1979). The vocal fold tissue was modeled as viscoelastic, transversely isotropic material. The material properties (shown in Table 4-1) were adopted from (Alipour, Berry and Titze, 2000) and (Xue *et al.*, 2012). Since vocal fold barely vibrates in the longitudinal direction, an in-plane motion constraint was implemented by employing relatively large values for the longitudinal Young's moduli (Cook and Mongeau, 2007). The boundary conditions for the vocal fold are shown in Figure 4.1 (b). A zero-displacement boundary condition was applied at the anterior, posterior and lateral surfaces, and a traction boundary condition was applied at the medial, inferior and superior surface. A simple hard-wall contact model was incorporated to model the collision of vocal folds. Two artificial non-slip and non-penetrable collision planes were placed one grid (in this case corresponds to ±0.1mm) off the medial plane to enforce a finite but small (0.2mm) minimum glottal gap.



(a) (b)

Figure 4.1 Simulation setup. (a) The computational domain and geometry of the vocal folds, larynx and vocal tract. (b) The inner-layer structure of the vocal fold as well as the boundary conditions applied on vocal folds walls.

Table 4-1 Material properties of the three inner layers of the vocal fold. ρ is the tissue density and η is the damping ratio.

Layer	Property						
	ρ (g/cm^3)	E_p (kPa)	v_p	E_{pz} (kPa)	v_{pz}	G_{pz} (kPa)	η (poise)
Cover	1.043	2.01	0.9	40	0.0	10	5
Ligament	1.043	3.31	0.9	66	0.0	40	7.5
Body	1.043	3.99	0.9	80	0.0	20	12.5

The entire geometry was immersed into a 2.4×10.8×14.5cm rectangular computational domain (Figure 4.1 (a)). For the flow model, a 0.8kPa pressure drop was applied between the inlet and outlet. A non-penetration non-slip boundary condition was applied at the vocal tract wall. The density of air was set as 1.146kg/m^3 at human body temperature. To alleviate computational cost, the kinematic viscosity of the air was set as 6.6×10^{-5}m^2/s, which corresponded to approximately 1/4 of the Reynolds number of normal human phonation. Such treatment would affect the turbulence flow in the supraglottal tract that is related with high frequency effects. For the acoustics model, a hard wall boundary condition was implemented on the vocal tract walls as $\frac{\partial \rho'}{\partial n} = 0, \frac{\partial p'}{\partial n} = 0, \vec{u}' \cdot \hat{n} = 0,$ where $\hat{n}$ is the outer normal vector. At the inlet, a zero Dirichlet boundary condition was applied as $p' = 0$ and $\vec{u}' = 0$. At the outlet, a complete reflection boundary condition was applied as $p' = 0$ and $\nabla \vec{u}' = 0$. The speed of sound was assumed to be 352 m/s.

Both the incompressible flow solver and acoustics solver employed a high resolution, non-uniform 64×256×192 Cartesian mesh, with highest grid density around the intraglottal region. The vocal fold was discretized by 28997 tetrahedral elements. The grid was based on our experience with previous three-dimensional simulations of similar configurations (Zheng *et al.*, 2010, 2011; Xue *et al.*, 2012, 2014; Xue and Zheng, 2017). A small time step of 1.149 μs was employed in the incompressible flow and solid

solvers, while 1/20 of this value was employed in the acoustics solver to provide a good temporal resolution as well as to satisfy the CFL stability constraint. The simulation was carried out 60000 steps on XSEDE COMET cluster, using 256 processors. The computational expense was about 15360 CPU hours per vibration cycle.

4.2 Results and discussion

4.2.1 Glottal flow waveform

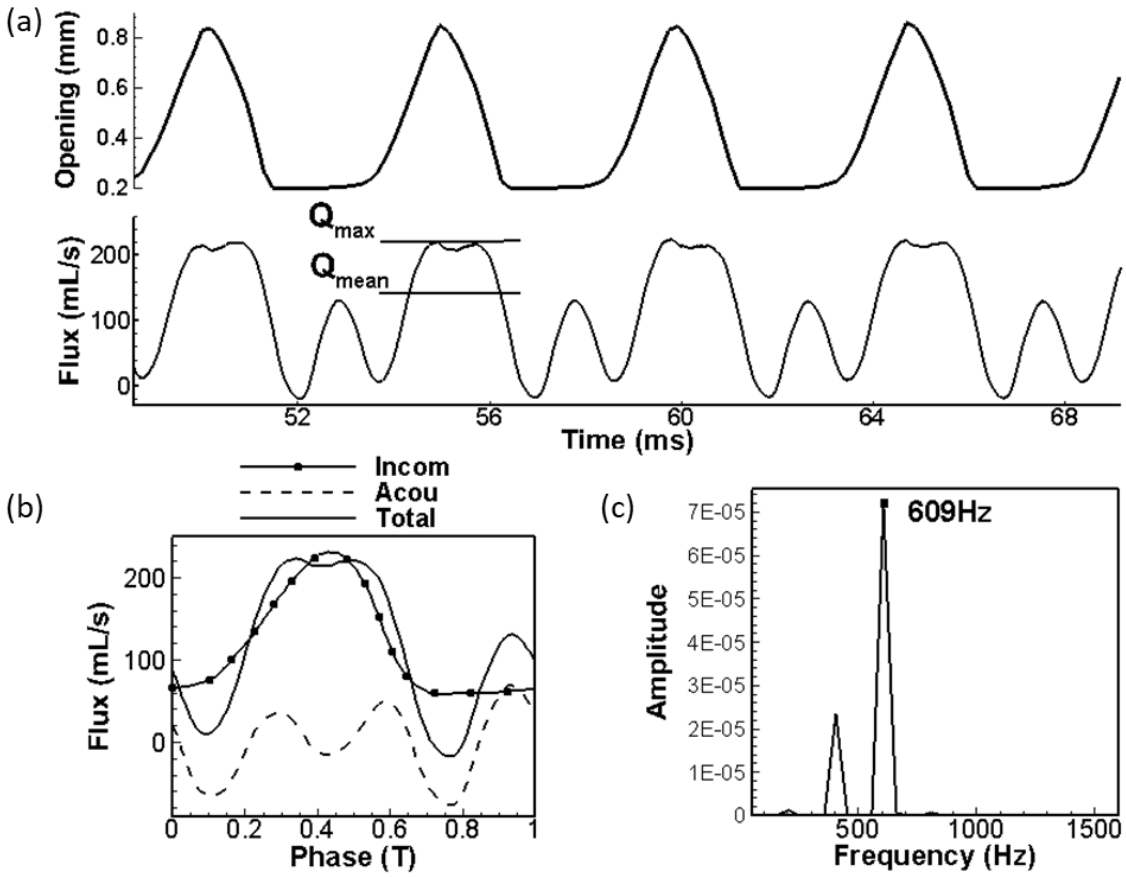


Figure 4.2 Glottal air flow rate. (a) Time history of glottal opening and air flow rate. (b) Phase averaged total flow, decomposed into incompressible part and acoustic perturbation part. (c) Spectrum analysis of acoustic flow rate.

The simulation was carried out for thirteen cycles, and the steady-state vibration was achieved at the 10th cycle. Figure 4.2(a) shows the time history of the glottal flow rate measured at the vocal tract outlet (mouth) as well as the opening size of the glottis during the last four cycles. The opening size of the glottis was calculated as the minimum distance between the two vocal folds at the mid-coronal plane. It shows that the shape of the flow rate generally followed the opening size. It increased with the increasing opening size and decreased with the decreasing opening size. It was noticed that during glottal closure (the opening size stayed at the minimum value), the flow rate did not stay at the minimum value but had a high peak. This high peak was mainly associated with the oscillation of the acoustic pressure associated with the first formant. The adopted hydrodynamic/acoustic splitting method allowed decomposing the total flow rate into the incompressible component and acoustic perturbation component, which are termed as the incompressible flow rate and acoustic flow rate, respectively, in the subsequent sections. Figure 4.2(b) shows the phase-averaged incompressible and acoustic flow rates along with the total flow rate. It shows that the incompressible flow rate is a typical glottal flow waveform with slow rise and rapid fall. The total flow rate generally followed the shape of the incompressible flow rate with fluctuations due to the oscillation of the acoustic flow rate. The strong peak during the glottal closure corresponded well with the peak of the acoustic flow rate at the same phase. Strong effects of acoustic coupling on glottal flow rate was also previously observed in the model of (Titze, 2006a, 2008) in which the glottal flow rate was calculated using the electrical circuit analogs method (Ishizaka and Flanagan, 1972a). Two to three evident ripples were generated on the flow rate waveform due to the acoustic coupling. In another work (Zañartu, Mongeau and Wodicka, 2007) which coupled the one-mass vocal fold model, Bernoulli flow model and wave reflection analog to simulate the fluid-structure-acoustics interaction, the depression of the flow rate during glottal opening was

observed, and this phenomenon was also shown to be due to the strong acoustic perturbation. It should be noticed that even though the glottis was closed, it was still possible to have high acoustic flow rate in the supraglottal tract due to the density change. It also needs to point out that the effect of acoustic coupling in the current model appeared stronger than in the aforementioned models. It was likely due to the total reflection boundary condition at the mouth and hard wall boundary condition at the vocal tract wall which had excluded the acoustic energy loss and would exaggerate the acoustic flow rate. However, such boundary conditions were considered as reasonable simplifications given small compliance of the vocal tract wall and very large area expansion at the mouth.

If the supraglottal tract is considered to be an ideal straight open-closed tube, its resonant frequencies can be analytically calculated as $F_n=(2n-1)c/4L$ (n=1, 2, 3 …) (Titze, 2000). For the current model with the length of 17.4cm and the speed of sound of 352m/s, the first three lowest formants of the tube would be 505Hz, 1517Hz, and 2528Hz. If the area variation along the tract was considered, these values would be shifted. (Story and Titze, 1998) have calculated the formants of the current supraglottal tract shape with a frequency domain transmission line technique (Sondhi and Schroeter, 1987), and they found the first and second formant as 628 Hz and 1510 Hz, respectively. The first formant has shifted significantly due to the area variation of the tract. Figure 4.2(c) shows the frequency spectrum of the acoustic flow rate obtained from the current simulation. Its dominant frequency was 609 Hz. This value was very close to the first formant calculated by (Story and Titze, 1998), indicating that the oscillation of the acoustic flow rate was dominated by the first formant resonance.

Several important voice quality-related parameters were computed based on the waveforms of the flow rate and opening size of the glottis. The average values as well as the physiological range of each quantity are listed in Table 4-2. It was found that these values were well within the physiological

range, indicating that the model had reproduced the essential biomechanics of voice production. It was noticed that the fundamental frequency (F_0) approached the higher end of the physiological range, which was likely due to the large values of longitudinal Young's moduli employed in the vocal fold model. It was of particular interest to look at the open quotient (τ_o) and the skewness quotient (τ_s). τ_o is defined as the duration of the open glottis divided by the period of the cycle, calculated based on the opening size of the glottis. τ_0 range from 0.4 to 0.7 for normal voice. A value lower than 0.4 indicates a "pressed" sound; a value above 0.7 indicates a "breathy" sound (Titze, 2000). In the current simulation, τ_o was 0.67, which was although within the normal range, indicated a breathier sound. τ_s is defined as the duration of flow acceleration divided by the duration of flow deceleration. This quantity was calculated based on the incompressible flow rate, as the flow acceleration and deceleration were more meaningful in the incompressible flow. τ_s typically ranges from 1.1 to 3.4 (Ishizaka and Flanagan, 1972b; LaMar, Qi and Xin, 2003; Duncan, Zhai and Scherer, 2006; Luo *et al.*, 2008). A higher value indicates that flow decreases rapidly in closing phase which will lead to higher vocal intensity (Baken and Orlinkoff, 2000; Titze, 2006b). In the current simulation, τ_s was 1.75, indicating moderate intensity. This was consistent with the 0.8kPa pressure drop across the larynx in the current simulation which was typical for moderate intensity voice production.

Table 4-2 Flow rate waveform related parameters.

	Computed value	Physiological range (Xue *et al.*, 2014)
F_0 (Hz)	203	60-250
Q_{mean} (mL/s)	143.43	110-220
Q_{max} (mL/s)	223.51	200-350
τ_o	0.67	0.4-0.7
τ_s	1.75	1.1-3.4

4.2.2 Acoustics

Figure 4.3(a) shows the frequency spectrum of the acoustic perturbation pressure (p') at the point of Y=4.0cm, Z=3.2cm, which was at the mid-coronal plane and 0.1cm above the vocal fold superior surface, indicated by the black dot in the inlaid graph. It shows that the energy of the third harmonic, which was 619Hz, and the seventh harmonics, which was 1444 Hz, has been boosted due to the close distance to the first and second formants of the tract shown in the previous section. The first formant of an open-close tube generates quart-wave resonance. To verify that, Figure 4.3 (b) shows the acoustic perturbation pressure (p') along the centerline of the vocal tract at different time instances during one vibration cycle. Distance represents the distance from the vocal fold superior surface, with positive value corresponding to the supraglottal tract and negative value corresponding to the subglottal tract. It shows that a standing wave was formed within the supraglottal tract. Acoustic perturbation pressure (p') oscillated in phase along the distance. The minimum vibration was at the supraglottal tract exit (distance=17.4cm) and maximum vibration was at the glottis exit (distance=0). This wave shape resembled the first mode of the quarter-wave resonator. Figure 4.3 (c)-(e) show the time-spatial variation of acoustic perturbation pressure (p'), incompressible pressure (p) and total pressure (P) along the centerline of vocal tract. The fluctuation of the wave amplitude of acoustic perturbation pressure (p') with time was noticed. It may be because of the highly turbulent flow pattern inside the vocal tract which generated fluctuation in pressures. It was interesting to see that while the frequency of the incompressible pressure (p) remained same as the vocal fold vibration frequency, the frequency of the total pressure (P) had been largely influenced by the acoustic perturbation pressure (p'), especially in the supraglottal tract region. This also suggested a strong acoustic coupling effect in the glottis.

Traditional linear source-filter theory of voice production assumes that the source-filter interaction was weak during normal phonation and so the acoustic pressure perturbation has little influence on the vocal fold vibration (Fant G., 1960; Flanagan, 1972). This assumption is made based on the fact that the vocal fold vibration frequency is normally well below the formants of the vocal tract so

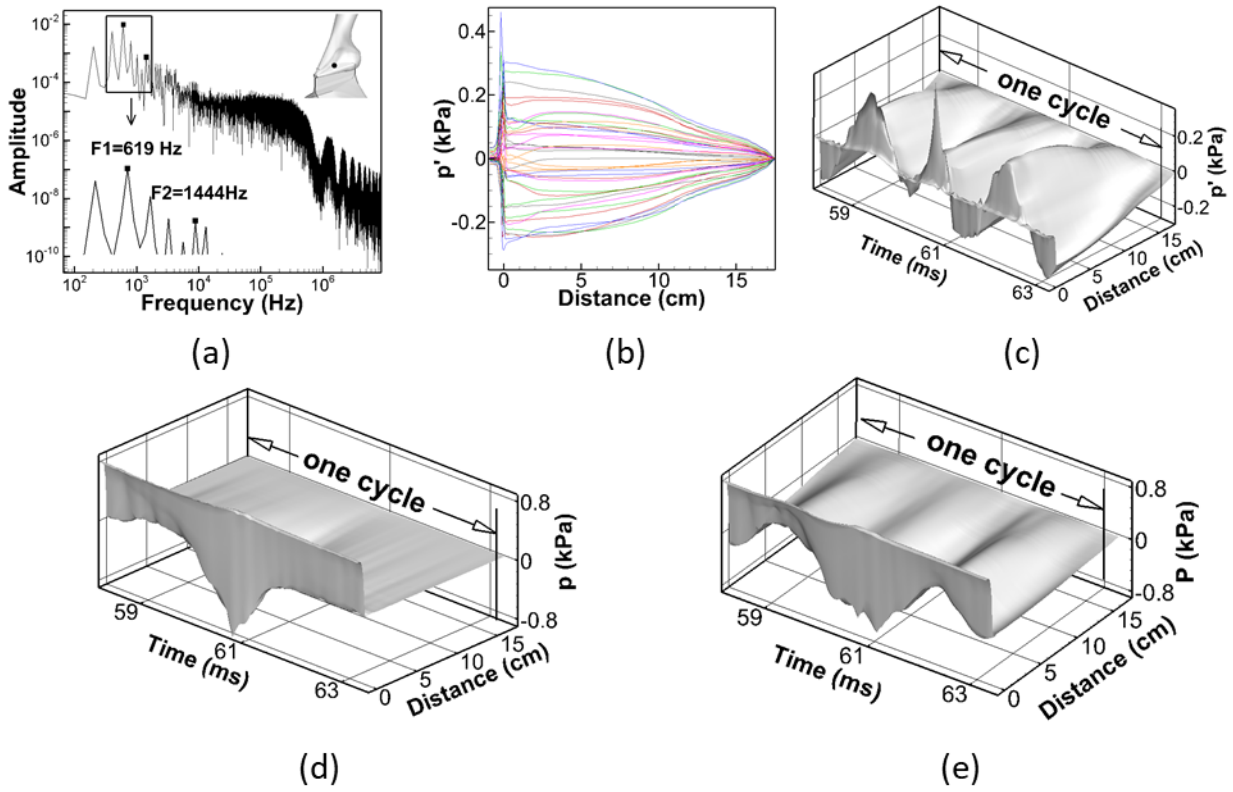


Figure 4.3 Pressure in the supraglottal tract. (a) Frequency spectrum of the acoustic perturbation pressure (p') at the point of Y=4.0cm, Z=3.2cm, which was at the mid-coronal plane and 0.1cm above the vocal fold superior surface, indicated by the black dot in the inlaid graph. (b) Acoustic perturbation pressure (p') along the centerline of the vocal tract at different time instants during one vibration cycle. (c)-(e): The time-spatial variation of acoustic perturbation pressure (p'), incompressible pressure (p) and total pressure (P) along the centerline of the vocal tract.

that the acoustic resonance does not happen. For this case, the acoustic pressure perturbation will be much smaller than the incompressible pressure so that it had little effect on vocal fold vibrations. However, the strong effect of acoustic perturbation pressure (p') on the total pressure (P) observed in Figure 4.3 (c) and Figure 4.2 (b) suggested that this assumption may not be valid. Figure 4.4 shows the time variation of the total pressure (P), acoustic perturbation pressure (p') and incompressible pressure (p) and at four different locations, three of which were within the glottis ((a) Y=3.6cm, (b) Y=3.7cm, and (c) Y=3.8cm) and one was just above the glottis ((d) Y=4.0cm). The variation of glottal opening is also plotted in the lower part of each subfigure. It shows that, firstly, throughout the cycle, acoustic perturbation pressure (p') and incompressible pressure (p) were at the same order at all positions. With the high value of acoustic perturbation pressure (p'), the driving force on the vocal folds as well as their vibrations will be significantly affected. Secondly, the total pressure (P) generally followed the shape of the incompressible flow pressure with fluctuations due to the oscillation of the acoustic perturbation pressure (p'). Thirdly, the effect of acoustic perturbation pressure (p') on the total pressure (P) was more significant toward the superior direction, suggesting stronger coupling effect inside the supraglottal tract. Therefore, these results suggested a strong acoustic-coupling effect during normal phonation, which maybe important to be included in future modeling works. Our observation was also supported by a recent study (Maxfield, Palaparthi and Titze, 2016) in which eight human subjects were recorded producing F_0 glide with arbitrary lengthened supraglottal vocal tract, even F_0 was well below the first formant of supraglottal tube, vocal fold vibration could be destabilized and resulted in F_0 jump.

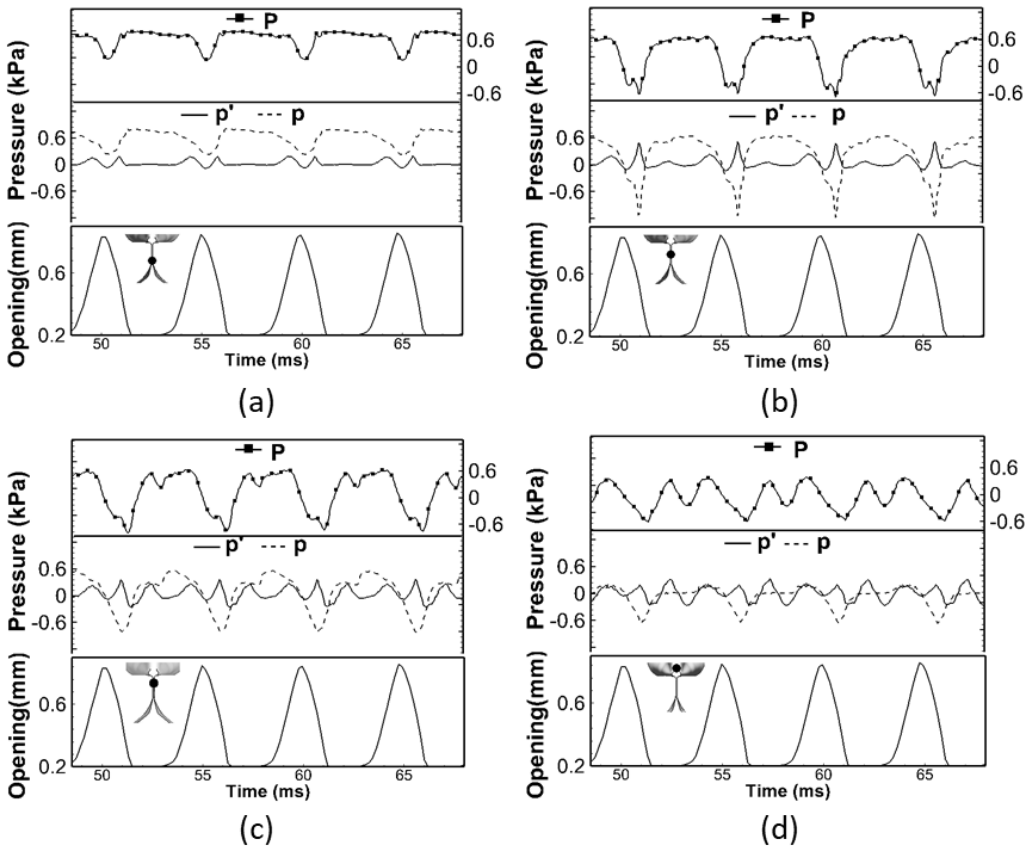


Figure 4.4 The time variation of the pressure at four different locations. (a) Y=3.6cm, (b) Y=3.7cm, (c) Y=3.8cm, (d) Y=4.0cm. P is total pressure, p' is acoustic perturbation pressure and p is incompressible pressure. The locations of the points are indicated by the black dot in the inlaid graph. The variation of glottal opening is also plotted in the lower part of each subfigure.

4.2.3 Vocal Fold Vibration

Figure 4.5 shows the vocal fold vibration pattern at four time instants during one vibration cycle. The first row shows the top view of the vocal folds, and the second row shows the vocal folds profile at the mid-coronal plane. It shows that the glottis presented a convergent shape during opening (instant a) and divergent shape during closing (instant c). At instant b, the glottis reached maximum opening and it

formed a straight channel. At instant d, the glottis was fully closed with the artificial gap remained. This convergent-divergent type of motion, also called the mucosal wave propagation, is an important indicator of healthy vocalization in clinic. From the mechanical point of view, it generates a temporal pressure asymmetry inside the glottis which ensures the net energy transfer from the airflow to vocal folds to sustain vibrations (Titze, 2000).

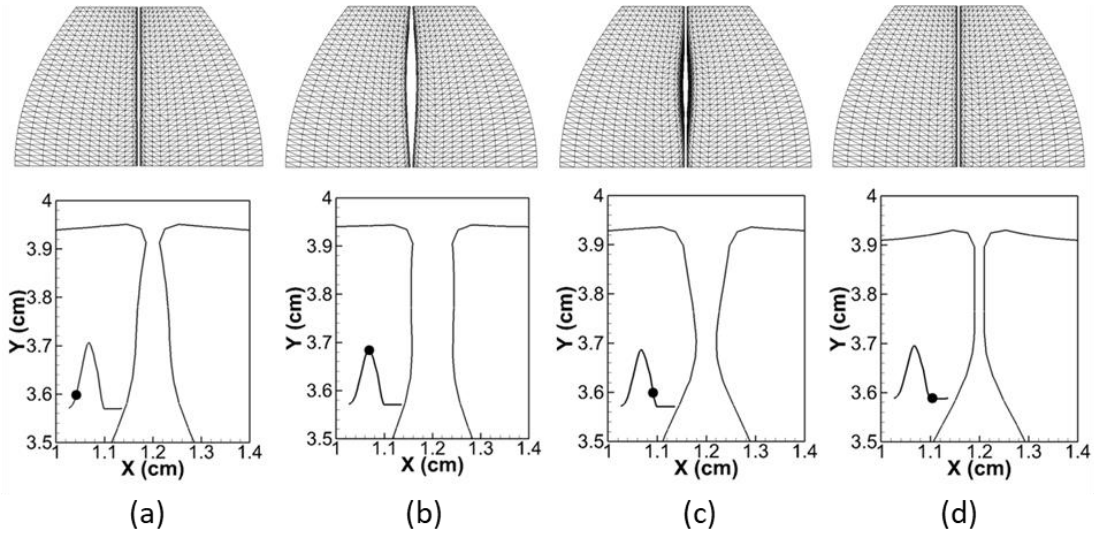


Figure 4.5 Vocal fold vibration pattern at four time instants during one vibration cycle. The first row shows the top view of the vocal folds, and the second row shows the profile of the vocal folds at the mid-coronal plane. The waveform of the opening size of the glottis is shown in the inlaid graph with the black dot superimposed indicating the time instant.

To have a quantitative analysis, the method of proper orthogonal decomposition (POD) was utilized to extract the dominant vibratory modes (Berry *et al.*, 1994; Zheng *et al.*, 2011). Figure 4.6 (a) shows the two most energetic modes. Mode 1 presented a clear convergent-divergent type motion, and Mode 2 was a lateral motion. The two modes captured 92% of the total vibration energy with Mode 1 and 2 was 70% and 22%, respectively. It should be pointed out that the shape of the two dominant

modes as well as the associated energy percentage were found to be very similar to (Berry *et al.*, 1994) which found that the first and second modes of vocal fold vibration in the simulation in (Alipour-Haghighi and Titze, 1985) captured 72% and 26% of the total energy, respectively. Figure 4.6 (b) is the time history of the modal coefficients of these two modes. The positive (negative) coefficient of Mode 1 corresponded to a convergent (divergent) shape, and the positive (negative) coefficient of Mode 2 corresponded to glottis abduction (adduction). The two coefficients oscillated with the frequency same as F_0, implying a 1:1 mode entrainment, which is an important indicator of normal phonation (Berry, 2001). Figure 4.6 (b) also shows that the increasing of the coefficient of Mode 2 was always companied by the positive coefficient of Mode 1, and vice versa. It indicated that the convergent glottal shape formed when vocal folds opened and a divergent shape formed when it closed, which was consistent with the observations in Figure 4.5.

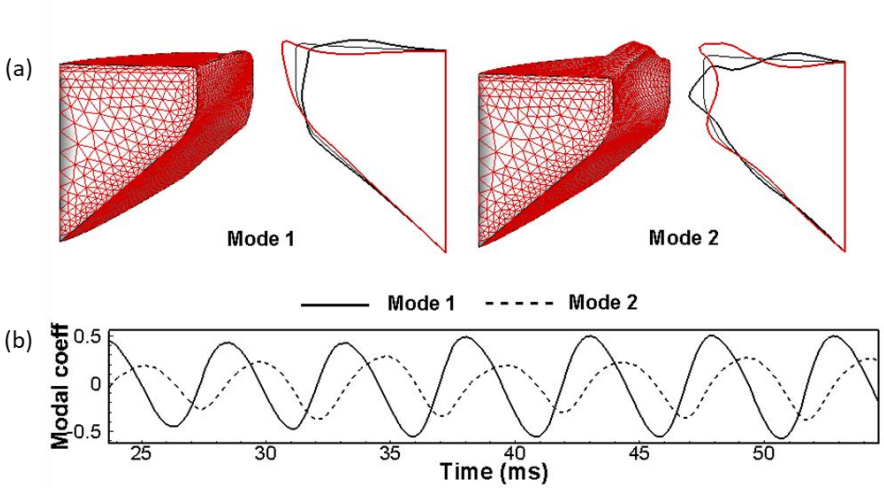


Figure 4.6 Empirical eigenmodes and the corresponding coefficients. (a) The three-dimensional and mid-coronal profile of the most energetic two empirical eigenmodes of the vocal fold at two-extreme phases through POD analysis. (b) Modal coefficients time history of the two eigenmodes.

4.3 Summary

A three-dimensional, first-principle based fluid-structure-acoustics interaction computer model of voice production which employed a more realistic human laryngeal and vocal tract geometries. Self-sustained vibrations and a reasonable glottal flow waveform were captured by the model and important voice-quality associated parameters were found to be well within the normal physiological ranges. The important convergent-divergent vibration pattern of vocal folds was captured. POD analysis demonstrated the 1:1 entrainment of the two dominant vibratory modes. The analogy between the vocal tract and a quarter-wave resonator was demonstrated. The simulation result reflected that the acoustic perturbed flux and pressure inside the glottis as well as the supraglottal tract were all at the same order with their incompressible counterparts, suggesting strong source-filter interactions during normal phonation.

The model demonstrated the capability of providing fully resolved and coupled fluid, structure and acoustics solutions in complex laryngeal shapes. It would be useful for studying the fundamental mechanisms of voice production, especially those related to the source-tract coupling effect, turbulent sound and different voice types. The model also can greatly extend the current framework of voice modeling to a wide range of pathological conditions which often involve complex vibration conditions.

The limitations of the current model have been listed. First, the current model assumed that vocal fold tissue was linear viscoelastic material. In general, vocal fold tissue exhibits a nonlinear stress-strain relationship (Min, Titze and Alipour-Haghighi, 1995; Chan and Titze, 1999; Zhang, Neubauer and Berry, 2007). However, this nonlinearity becomes obvious only during large deformation events such as posturing. During phonation, vocal folds exhibit a nearly linear stress strain relationship when active muscular tension is present (Titze, 2006a). Therefore, the material properties adopted in the current

study can be interpreted as properties for a given posturing. Second, the current model had reduced the Reynolds number to 1/4 of the normal value to alleviate the computational cost. Such treatment would affect the turbulence flow in the vocal tract which is related to high frequency effects. Third, an artificial gap was imposed between the two vocal folds, allowing leakage flow even during what would be considered as glottal closure. Fourth, the current model had employed the total reflection boundary condition at the mouth and hard wall boundary condition at the vocal tract wall which had excluded the acoustic energy loss and would exaggerate the acoustic flux. However, such boundary conditions were considered as reasonable simplifications given small compliance of the vocal tract wall and very large area expansion at the mouth. Last, the calculation of Navier-Stokes equation based fluid solver requires more computational resource and time than the reduced order based solver, limiting the ability to perform real-time simulation. However, the continually increasing computer ability and the parallel computing technique would decrease the computational time in the future.

CHAPTER 5 EFFECT OF LONGITUDINAL VARIATION OF LAYER THICKNESS

In this chapter, the influence of longitudinal variation of the vocal fold inner layer thickness has been studied. This work is reproduced from "Jiang, Weili, Qian Xue, and Xudong Zheng. 'Effect of longitudinal variation of vocal fold inner layer thickness on fluid-structure interaction during voice production.' Journal of biomechanical engineering 140, no. 12 (2018)."

5.1 Introduction

In the human vocal fold, the thickness of each layer is known to vary along the longitudinal direction (Hirano, Kurita and Nakashima, 1981). Usually, the stiffer ligament layer is thicker at the sites of attachment at the anterior commissure and the vocal process of the arytenoid to withstand the massive stress of vocal fold action. The softer cover layer is thicker at the middle to provide a cushion for vocal fold collision (Titze, 2000). Figure 5.1 shows the averaged thickness of the cover and ligament layers at different longitudinal positions of five male vocal fold samples measured by (Hirano, Kurita and Nakashima, 1981). The square and diamond symbols represent the original measurements of the ligament and cover, respectively, and the curves are the fourth-order polynomial curves fitting the data. According to the data, the variation of the thickness in the longitudinal direction is up to 57% and 43% for the cover and ligament, respectively, about their average values. These are relatively large variations which could have a significant impact on stiffness distribution on the vocal fold. Based upon the aforementioned findings that inner layer stiffness can significantly affect vocal fold vibration amplitude and pattern, the longitudinal variation of thickness might affect vocal fold vibration through its effect on stiffness. (Bhattacharya, Kelleher and Siegmund, 2015) has computationally investigated the influence of longitudinal stiffness variation of the vocal fold on vibration by changing the elastic modulus. It was found that the stiffness variation could affect the fundamental frequency and flow rate and caused

anterior-posterior asymmetry and a wave motion in the longitudinal direction. However, in that study, the vocal fold was modeled as a one-layer isotropic material. Vocal fold is anisotropic and it was reported that the anisotropy could significantly suppress the wave motion in the longitudinal direction (Zhang, 2014). Moreover, vocal fold is pre-stressed in the longitudinal direction during posturing. This condition will stiffen the vocal fold and further suppress the wave motion in the longitudinal direction. However, this feature was not modeled in (Bhattacharya, Kelleher and Siegmund, 2015). Therefore, the model may have over-predicted the wave motion in the longitudinal direction. In fact, it was observed in the experiment that vocal fold vibration was mostly in the transverse plane and very little in the longitudinal direction(Berry, Montequin and Tayama, 2001; Cook, Nauman and Mongeau, 2008). Therefore, it remains unclear how the stiffness variation in the longitudinal direction affects vocal fold vibrations if the anisotropy and longitudinal stiffening was considered. Furthermore, it is also unknown to what extent the thickness variation in inner layers affects the stiffness variation on vocal fold. Such understanding would be important to estimate the ultimate effect of thickness variation on vocal fold vibrations. (Xue *et al.*, 2011) utilized an anisotropic vocal fold model to investigate the effect of longitudinal variation of thickness on vocal fold eigenmodes and found that the eigenmodes were insensitive to such variations; yet, it is unclear whether this conclusion still holds when fluid-structure interactions is considered.

Therefore, in the present chapter, numerical approach validated in the previous chapter was utilized to investigate the effect of longitudinal variation of inner layer thickness on glottal flow dynamics and vocal fold vibrations. However, the acoustic solver was excluded as our focus is investigating the effect on the vocal fold dynamics. Four vocal fold models of different degrees of variations were generated for the purpose of covering the limit of inter-subject variation. The effect of

thickness variation on vocal fold stiffness distribution, glottal flow waveform, vocal fold vibration pattern, glottal angle, and energy transfer was studied by comparing between the cases. The possible underlying mechanism of the differences was also discussed.

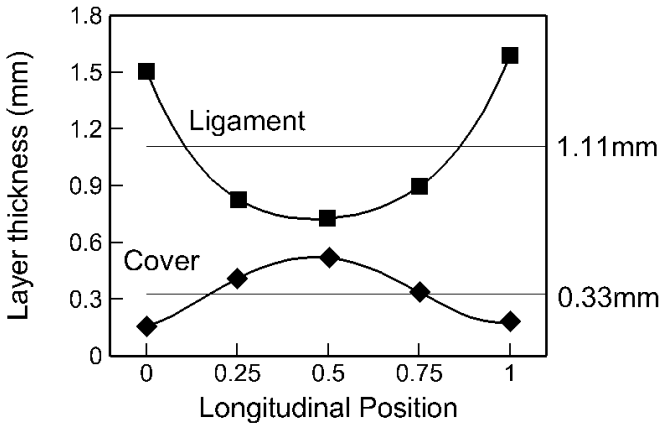


Figure 5.1 The averaged thickness of the cover and ligament layers of five male vocal fold samples measured. The square and diamond symbols represent the original measurement of ligament layer and cover layer, respectively. Fourth-order polynomial curves are applied to fit the data: $y=5.92x^4-12.16x^3+11.16x^2-4.84x+1.51$ for the ligament and $y=5.10x^4-9.29x^3+3.60x^2+0.62x+0.15$ for the cover, where x is the non-dimensioned longitudinal position and y is the layer thickness (mm). 1.11mm and 0.33mm are the averaged layer thicknesses for the ligament and cover, respectively.

5.2 Simulation setup

The vocal fold inner layer thickness was modified based on the model in Chapter 4. Four vocal fold models with different levels of variation of layer thickness were created. The first model was a uniform thickness model. The thickness of each layer was based on the average value of Hirano's measurement, as shown in Figure 5.1, which was 0.33mm for the cover and 1.11mm for the ligament. In the second model, the thickness variation was directly based on Hirano's data by using the fitting curves

as shown in Figure 5.1. In the context below, these two cases are denoted as the baseline model and Hirano model, respectively.

Nevertheless, Hirano's measurement was the average value of five subjects (Hirano, Kurita and Nakashima, 1981). To include the possible extreme conditions of inter-subject variation, we created two

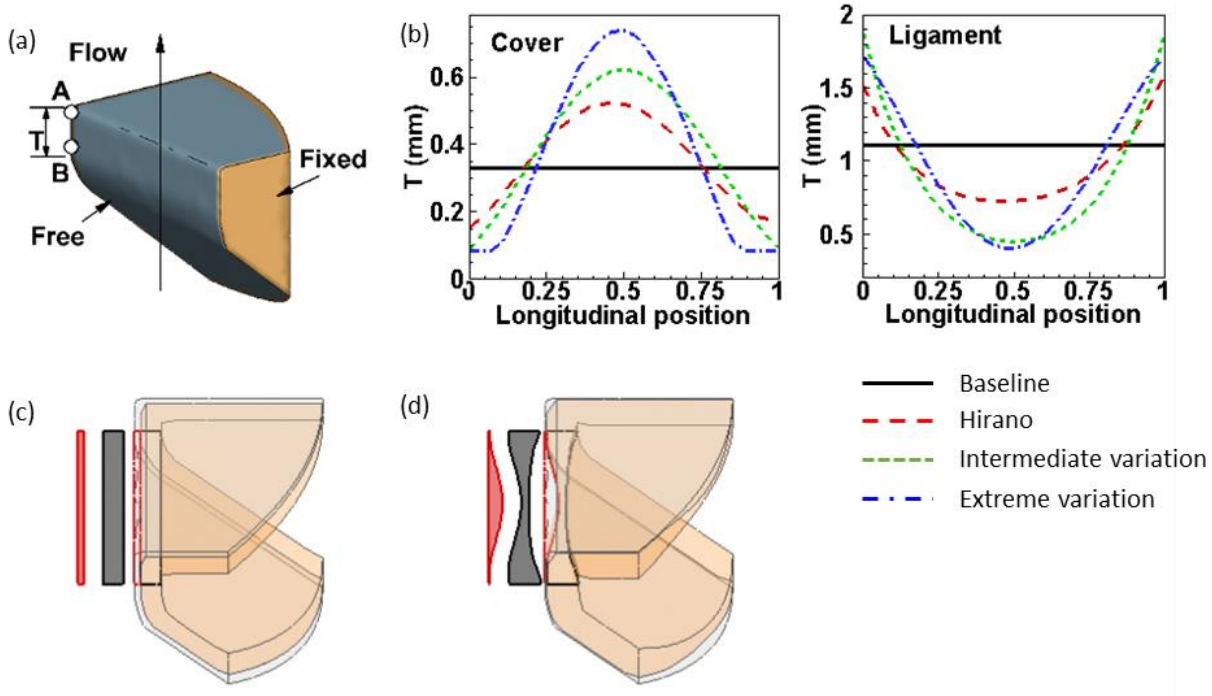


Figure 5.2 The longitudinal thickness variation in the four parametric cases. (a) The geometry and boundary conditions of the vocal fold. T denotes the height of the medial surface. Points A and B locate at Y=3.86cm and Y=3.66cm, respectively, which were used in the calculation of the glottal angles; (b) The varied thickness along the longitudinal direction of the cover and ligament layers in the four cases; (c) &(d) The three-dimensional configurations of the cover and ligament layers of the vocal fold in the baseline case and extreme variation case, respectively. The shapes of the cross section in the vertical direction in the medial surface of the cover and ligament layers were plotted beside the vocal fold with the color of red and black, respectively.

more cases to have more significant changes of the thickness in the longitudinal direction. The methods for creating the models are briefly described below. A thickness ratio, TR, was defined for each layer, which was computed as the ratio of the largest thickness to smallest thickness along the longitudinal direction. TR_c and TR_l denote the ratio of the cover and ligament, respectively. In the Hirano model, TR_c =3.38 and TR_l =2.07. In the third model, we doubled the values of TR to be TR_c =6.77 and TR_l =4.15. In the fourth model, the values of TR were further increased to be TR_c =8.67 and TR_l =4.28. These two values were chosen because further increasing TR_c and TR_l only resulted in small changes in the largest thickness of each layer. For example, further doubling TR_c and TR_l beyond this point only resulted in less than 10% changes in the largest thickness in the cover and ligament layers. Therefore, the fourth model represented an extreme condition. In the context below, the third and fourth models are denoted as the intermediate variation case and extreme variation case, respectively. The absolute values of the thickness at the five longitudinal positions as in the Hirano's measurement (referred to Figure 5.1) were computed by using the TR_c and TR_l and based on the constraint that the average value remained the same and the assumption that the thickness at the anterior and posterior ends were equal. Figure 5.2 (b) shows the variation of the thickness of the cover and ligament layers in the four models, respectively.

To clearly illustrate the internal structure, Figure 5.2 (c) and (d) show the three-dimensional configurations of the inner layers of the baseline model and the extreme variation model, respectively. The shapes of the cross section in the vertical direction of the medial surface of the cover and ligament layers were plotted beside the vocal folds in red and black, respectively. It shows that in the extreme variation model, the cover layer is thickest at the middle and the ligament layer is thickest at the anterior and posterior ends.

A 1.0 kPa pressure drop was applied between the outlet and inlet of the vocal tract. Non-slip and non-penetration boundary conditions were applied on the vocal tract walls. The kinematic viscosity of the air was set to be 1.65×10^{-5} m^2/s corresponding to the human body temperature of 37°C. A zero-displacement boundary condition was applied at the anterior, posterior and lateral surfaces of the vocal fold and other surfaces were free to move. The simulations were carried out on the XSEDE COMET cluster, using 128 processors. The computational expense was about 4600 CPU hours per vibration cycle which equals to 1.5 days per cycle.

5.3 Results and discussion

5.3.1 Effect on stiffness distribution on vocal fold surface

Because the ligament is stiffer than the cover, the spatial variation of layer thickness would affect the stiffness distribution on the vocal fold. Specifically, the thickened cover and thinned ligament at the middle of the vocal fold would reduce the stiffness at the position, and the thinned cover and thickened ligament at the anterior/posterior aspects would increase the stiffness at the position. To examine the extent to which the thickness variation affected the stiffness distribution, a numerical indentation technique was used to measure the stiffness on the vocal fold surface. The indentation was implemented by applying a uniform pressure loading on a small area and then solving the deformation using the finite element solver. The stiffness, defined as the slope of the force-displacement curve, can be calculated as $k=PA/\Delta x$, where k is the stiffness, P is the pressure loading, A is the area and Δx is the maximum deformation. The details about the numerical indentation method can be found in (Geng, Xue and Zheng, 2017). To obtain the distribution of the stiffness, the medial surface of the vocal fold was evenly divided into 21 sections with 7 sections in the longitudinal direction and 3 sections in the vertical direction (shown in Figure 5.3). The portion that is close to the two ends was excluded since they were

constrained by the fixed boundary condition and would barely vibrate. The numerical indentation was applied at each section. Figure 5.3 shows the stiffness distribution along the longitudinal direction on the medial surface of the four vocal fold models by averaging the values in the vertical direction. It reflects that, for all the vocal fold models, the stiffness was lowest at the center and highest at the posterior end. It needs to be pointed out that stiffness is a function of many properties including the material properties, material orientation, geometric dimensions, type of constraint, loading direction and position. Therefore, even though the layers were uniform in the baseline model, the stiffness was still increased toward the anterior/posterior end and that was the effect of fixed boundary condition. Generally, a larger thickness variation resulted in a larger stiffness variation. Such effect was most prominent when comparing the baseline model and Hirano model. The thickness variation has resulted in a much lower stiffness in the middle and higher stiffness toward the anterior/posterior ends. It was also interesting to notice that further increasing the thickness variation in the intermediate variation model and extreme variation model only resulted in a small change in the stiffness at most places of the vocal fold except near the anterior/posterior ends. That means the stiffness distribution became insensitive to the thickness variation once the thickness variation level exceeded that in the Hirano model. Quantitatively, by using the values in the baseline model as the normalized values, the stiffness decreased by 16%, 21%, 24% at the middle, increased by 18%, 12%, 47% near the anterior end, and increased by 5%, 10%, 29% near the posterior end for the Hirano model, intermediate variation model and extreme variation model, respectively. The average stiffness (k_{ave} in Table 5-1) over the entire medial surface was computed by averaging the stiffness values in the 21 sections. Interestingly, it was found that the value only changed by -3%, -5%, +5% in the Hirano model, intermediate variation model, and extreme variation model, respectively. Considering that the same average thickness was kept for

each layer among the cases, it implied that the average stiffness was highly correlated with the average thickness of the inner layers.

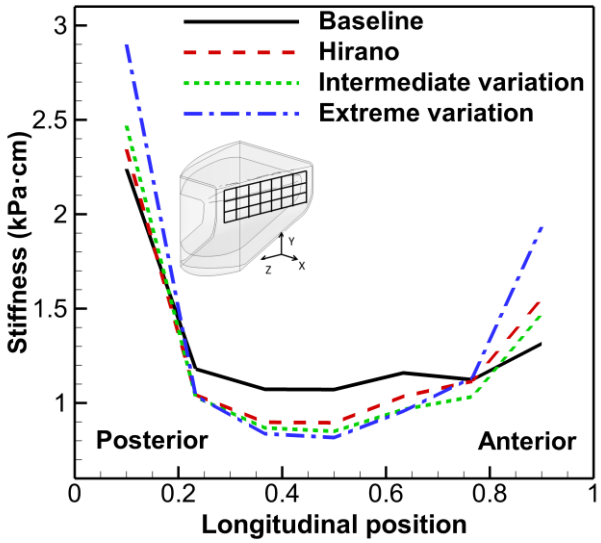


Figure 5.3 The stiffness distributions in the longitudinal direction at the medial surface of the four vocal fold models. The subfigure illustrates the sections used in the numerical indentation.

5.3.2 Effect on the waveforms of the glottal flowrate and opening size

The time history plot of the glottal flow rate and glottal opening size of the baseline case was shown in Figure 5.4 (a). And the phase-averaged flow rate and opening size for the four cases was shown in Figure 5.4 (b). The opening size was defined as the minimum distance between the two vocal folds across the vertical direction in the mid-coronal plane. Based on the waveforms of the flow rate and opening size, several voice-quality related quantities were calculated for the four cases and listed in Table 5-1. These include the fundamental frequency (F_0), open quotient (τ_o), skewness quotient (τ_s), maximum flow declination rate (MFDR), maximum flow rate (Q_{max}), mean flow rate (Q_{mean}) and maximum glottal opening (Gap_{max}). τ_o and τ_s were calculated based on the waveform of the glottal flow

rate. MFDR was obtained by calculating the time derivative of the flow rate during the closing phase of the glottis. It should be pointed out that these quantities were computed by using the phase-averaged flow rate and opening size which has eliminated the effect of cyclic variation.

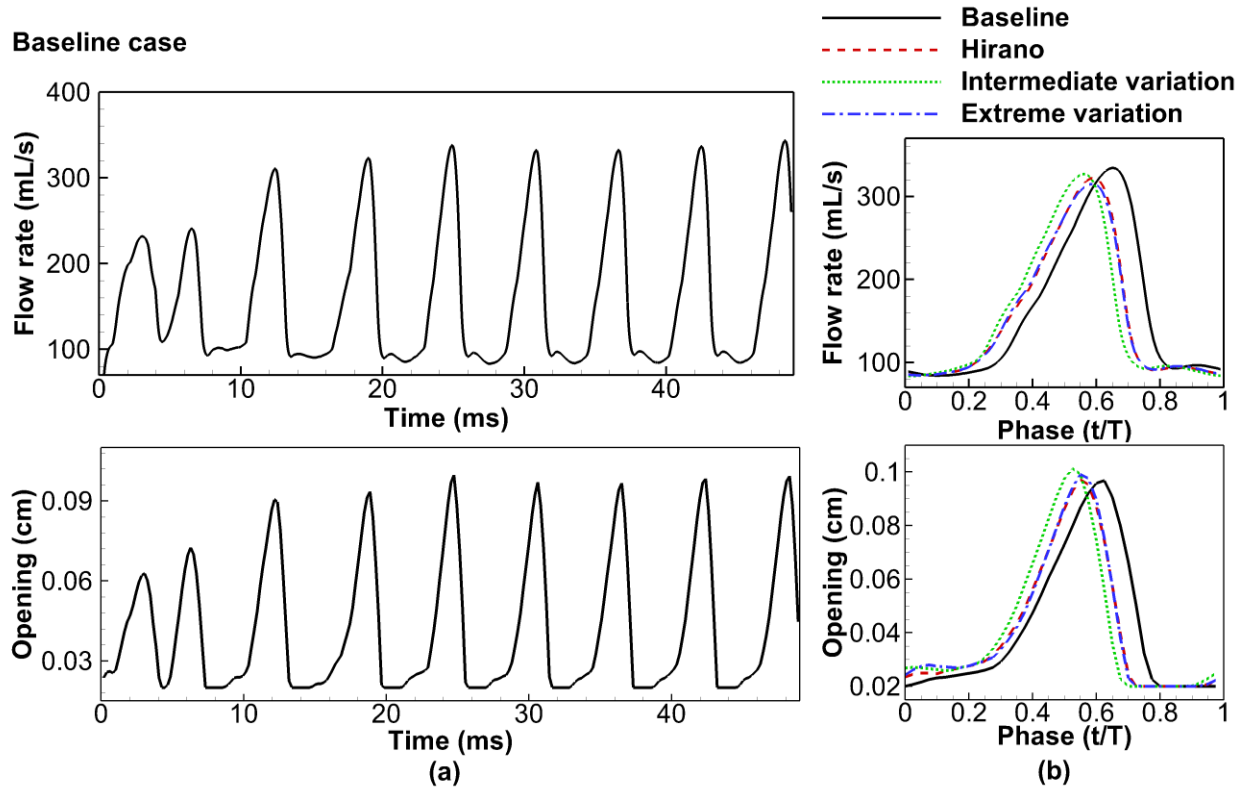


Figure 5.4 The glottal flow rate and opening for the four cases. (a) The time history of the glottal flow rate and glottal opening of the baseline case. (b) The phase-averaged flow rate and glottal opening for the four cases.

To compare between the cases, the relative differences of the parameters by using the values in the baseline case as the normalized values were computed. They are shown in the round brackets in Table 2. Moreover, the percent values of cyclic variation of some of the parameters, which was computed as the root-mean-square of the variation of each cycle from the mean value, were also

computed and listed in the square brackets. It was found that the differences with the baseline case in these parameters were generally below 5% except that τ_s in the Hirano model increased by 8% and Q_{max} in the extreme variation case decreased by 5.9%. The cyclic variations were generally at the same level. Therefore, the thickness variations did not have a significant effect on the glottal flow and opening size waveforms.

It is generally assumed that, with the reduced stiffness at the middle of the vocal fold, the vibration amplitude and flow rate would increase because the middle area is where the largest vibration occurs. However, in our simulations, although the stiffness was reduced by 16%-24% at the middle in the cases, the vibration amplitude and flow rate remained nearly the same as the baseline case. Considering that the average stiffness remained nearly the same among the cases, our simulation results implied that the vibration amplitude and flow rate were determined by the average stiffness, not the local stiffness. This result contradicted with the conclusion of a previous study which applied modulus gradients in the longitudinal direction in an isotropic vocal fold model and found that the stiffness gradients affected the fundamental frequency and flow rate (Bhattacharya, Kelleher and Siegmund, 2015). However, as aforementioned in the introduction, the model employed the isotropic material assumption and did not incorporate the effect of longitudinal stiffening of vocal fold, and therefore created a strong wave motion in the longitudinal direction. This artificial wave motion would affect the motion in the transverse plane and further affect the fundamental frequency and flow rate. Furthermore, it was unclear if the average stiffness was maintained among the cases in (Bhattacharya, Kelleher and Siegmund, 2015). Therefore, it was unclear if the observed effect due to the stiffness gradient or the change of the average stiffness.

Table 5-1 The glottal flow rate and opening waveforms related parameters. The numbers in the round brackets denote the difference with the baseline value (%). The numbers in the square brackets denote the percent value of root-mean-square cyclic oscillation (%).

	Baseline case	Hirano case	Intermediate variation case	Extreme variation case	Physiological range (Xue *et al.*, 2012)
F_0 (Hz)	170.3 [0.8]	167.9 [0.9] (-1.34)	165.8 [0.9] (-2.61)	168.7 [1.0] (-0.89)	60-250
τ_o	0.537 [1.1]	0.539 [1.3] (0.26)	0.535 [1.6] (-0.36)	0.544 [1.0] (1.16)	0.4-0.7
τ_s	2.29 [2.5]	2.48 [3.9] (8.05)	2.30 [3.0] (0.37)	2.38 [5.3] (3.94)	1.1-3
MFDR (m^3/s^2)	0.479 [2.7]	0.468 [4.9] (-2.12)	0.473 [2.0] (-1.12)	0.462 [3.5] (-3.41)	-
Q_{mean} (mL/s)	159.8 [1.6]	154.7 [0.5] (-3.2)	154.7 [0.8] (-3.2)	153.8 [1.5] (-3.7)	110-220
Q_{max} (mL/s)	333.9 [0.8]	323.1 [2.8] (-3.5)	326.9 [0.7] (-2.4)	315.1 [2.6] (-5.9)	200-350
Gap_{max} (mm)	0.98 [1.2]	0.978 [5.3] (-0.2)	1.023 [1.8] (4.4)	1.001 [3.3] (2.14)	-
k_{ave} (kPa·cm)	1.31	1.27 (-3.0)	1.24 (-5.1)	1.37 (5.0)	-

5.3.3 Effect on vocal fold vibration pattern

While the four cases showed nearly the same vibration amplitude, it would be of interest to examine their vibration patterns. When the vocal folds vibrate, a mucosal wave forms and propagates in the cover layer. This wave generates an inferior-superior phase difference in the motion, creating an alternating convergent-divergent glottal shape, which is a key mechanism for sustained energy transfer

from the air flow to vocal folds. Figure 5.5 plots the mid-coronal profile of the vocal folds at four time instants during one vibration cycle for the baseline case and the extreme variation case. The time instant associated with each subfigure is indicated as a black dot on the flow rate plot shown at the right-bottom. It shows that the glottis was alternating convergent-divergent shapes in both cases. At instant (a), the glottis was about to open with a very clear convergent shape in the inferior-superior direction. At instant (b), the glottis reached its maximum opening, forming a nearly straight channel. Due to flow inertia, the glottal flow was still increasing although the glottis had already started to close. At instant (c), the glottis was closing with a very clear divergent shape, and the glottal flow was decreasing. At instant (d), the glottis was fully closed (with artificial gap). It is also important to see that the profiles from the two cases nearly overlapped each other, indicating the high similarity of the vibration pattern.

To have a quantitative comparison of the vibration pattern between the cases, the technique of proper orthogonal decomposition (POD) was applied to extract the principle modes of the vibration (Berry *et al.*, 1994; Zheng *et al.*, 2011). As the vibration was nearly symmetric about the centerline, only the left vocal fold was analyzed. Figure 5.6 (a) and (b) show the three-dimensional shape as well as the mid-coronal profile of the vocal fold of the two most energetic modes of the baseline case. The two modes together captured 93% of the total kinetic energy, 82% and 11% in each mode, respectively. Mode 1 was mostly an in-phase lateral motion and Mode 2 was an out-of-phase lateral motion. The same modal shapes were observed in other cases. The energy associated with the two modes was 80% and 12% in the Hirano model, 77% and 13% in intermediate variation case, 78% and 13% in extreme variation case, respectively. To precisely quantify the similarity of the modal shapes, the dot-products of the unit vectors of the two corresponding modes between the baseline case and the other three cases were computed. The value of one indicated an exact match of the modal shape, and value of zero

indicated complete orthogonality. The results showed that the dot-products of the two corresponding modes between the baseline case and other cases were all above 0.98, indicating nearly the same vibration patterns between the cases.

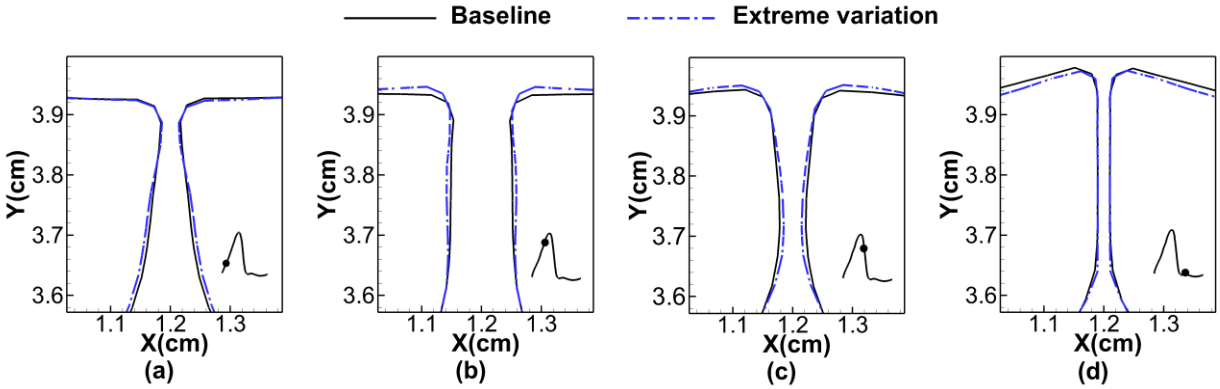


Figure 5.5 The mid-coronal profiles of the vocal folds at four time instants during one vibration cycle in the baseline and extreme variation cases. The corresponding time instant is indicated as a black dot on the flow rate plot shown at the right-bottom of each subfigure.

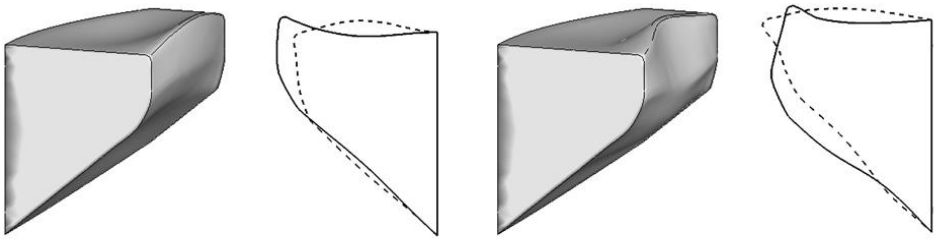


Figure 5.6 The two most energetic modes of the baseline case. (a) Mode 1; (b) Mode 2. In each subfigure, the left side is the three-dimensional shape and the right side is the mid-coronal profile of the vocal fold.

5.3.4 Effect on glottal angles

Figure 5.7 shows the phase-averaged variation of the glottal angle at three coronal planes for the four cases. Because the vocal folds vibrated symmetrically about the mid-sagittal plane, the glottal

angle was measured between the right vocal fold and mid-sagittal plane. It was calculated by assuming a straight line between two points on the vocal fold that were initially at Y=3.86 and Y=3.66cm (indicated in Figure 5.2(a)). A positive angle corresponded to a convergent shape and a negative angle corresponded to a divergent shape. Two common features were observed for all the cases. First, the glottal angle decreased from the mid-coronal plane to the anterior and posterior ends. It is easy to understand because the vocal folds were fixed at the anterior and posterior ends so that the vibration was gradually constrained toward them. Second, at each plane, the maximum convergent angle was larger than the maximum divergent angle (absolute value). At the mid-coronal plane (Figure 5.7 (a)), the maximum convergent angles were 6.57±0.11°, 7.62±0.50°, 8.45±0.25°, 8.79±0.21° in the baseline, Hirano, intermediate variation, extreme variation cases, respectively, and the maximum divergent angles were 3.05±0.16°, 4.03±0.52°, 4.74±0.36°, 4.24±0.81°, respectively. The second number denotes the root-mean-square value of the cyclic variation. This feature may be related to the slow opening and fast closing of the vocal folds so that there was less time for developing the angle during closing.

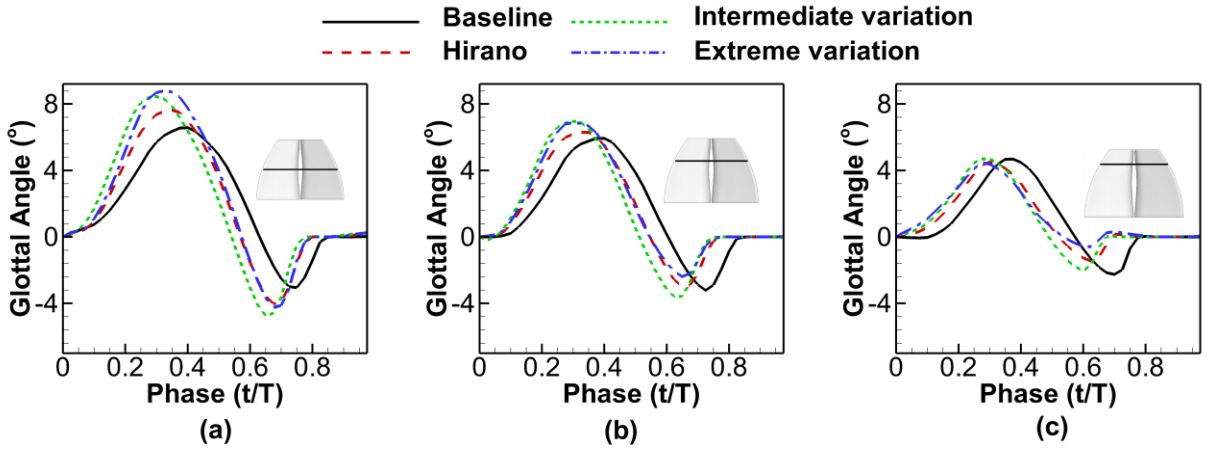


Figure 5.7 The phase-averaged glottal angle variation at the three coronal planes. The position of each coronal plane is denoted at the top-right corner of each figure.

Differences between these cases were also observed. In Figure 5.7 (a) which is for the mid-coronal plane, the baseline case had the smallest maximum angles in both the convergent and divergent shapes. A higher level of thickness variation generated larger maximum angles. However, the difference between the cases gradually vanished toward the anterior end, as shown in Figure 5.7 (b) to Figure 5.7 (c). To further illustrate this difference, Figure 5.8 (a) and (b) show the maximum convergent and divergent angles at different longitudinal positions for the cases. It was found that the thickness variations generated larger maximum convergent and divergent angles in the middle of the vocal fold and smaller angles near the anterior and posterior ends. Specifically, Figure 8(a) shows that the thickness variations had a large effect on the convergent angles at the middle of the vocal fold and small effect near the anterior and posterior ends. The effect at the middle of the vocal fold was monotonic that a higher level of thickness variation generated larger angles. Interestingly, Figure 5.8 (b) shows a non-monotonic trend that the intermediate variation case generated the largest maximum divergent angle in the middle while the extreme variation case generated the smallest maximum divergent angle near the anterior and posterior ends. Also, the intermediate variation case had a larger maximum divergent angle than the Hirano model near the anterior end. A closer comparison with the stiffness values in Figure 5.3 revealed that this non-monotonic trend was highly correlated to the stiffness distributions. It was observed in Figure 5.3 that the intermediate variation case had the smallest stiffness at the longitudinal position between 0.65 - 0.8 and smaller stiffness than the Hirano model near the anterior end. It was consistent with the trend of divergent angles at these positions. It was noticed that the intermediate variation case did not have the smallest stiffness at the middle but had the largest divergent angles there. It could be the effect of the adjacent tissues at the longitudinal position between 0.65 - 0.8 which had much larger angles than other cases. The above observations implied that the

glottal angles were largely affected by the local stiffness. Such relationship was more prominent in the divergent angles which might be due to that glottal closing is primarily driven by the elastic recoil force inside the vocal folds which are directly related to the stiffness.

It needs to be pointed out that the maximum convergent angles increased by up to 34% and the maximum divergent angles increased by up to 55% among the thickness variation cases by comparing to the baseline case. However, the absolute changes were small which was less than 2.3° for the convergent angle and 1.8° for the divergent angle. Considering that the angle was calculated for half of the glottis, the absolute change would be less than 4.6° and 3.6° for convergent and divergent angles, respectively, if the full angle between the two vocal folds was used. This small change in the glottal angle did not result in noticeable difference in the vibration patterns.

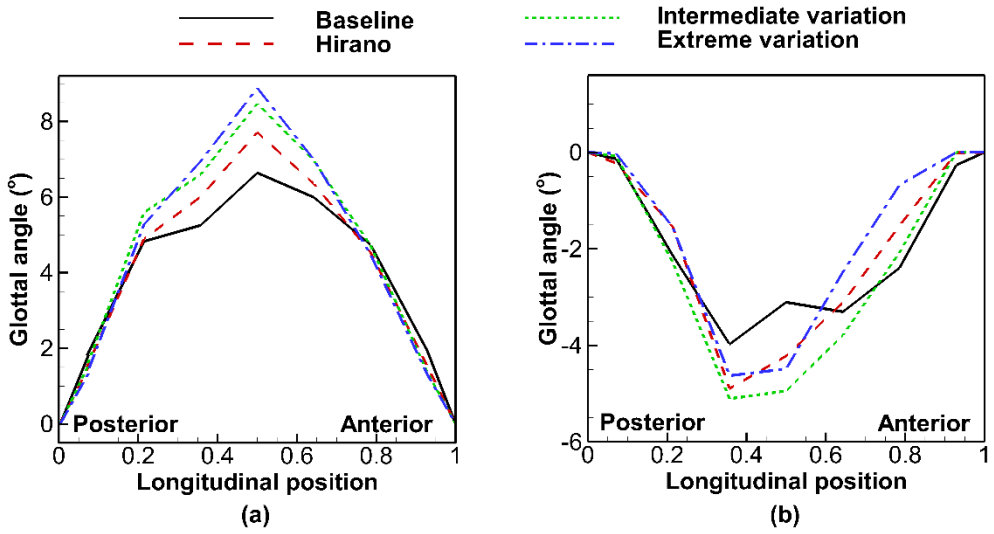


Figure 5.8 The phase-averaged maximum glottal angle distribution in the longitudinal direction. (a) convergent angle and (b) divergent angle.

5.3.5 Effect on energy transfer

Figure 5.9 (a) shows the phase-averaged power transfer from the air flow to the vocal fold during one vibration cycle for the four cases. Figure 5.9 (b) shows the phase-averaged glottal opening size. The power was calculated by multiplying the flow pressure with the normal components of the velocity vectors and integrating over the entire vocal fold surface. The contribution of the viscous force to the power was neglected since it is usually two orders of magnitude smaller than the pressure contribution (Thomson, Mongeau and Frankel, 2005). It was interesting to see that, for all these cases, the power was positive at the very early opening phase and most of the closing phase. During other times, it was negative. It suggested that the aerodynamic pressure facilitated the vocal fold motion during the early opening phase and most of the closing phase and retarded its motion at other times.

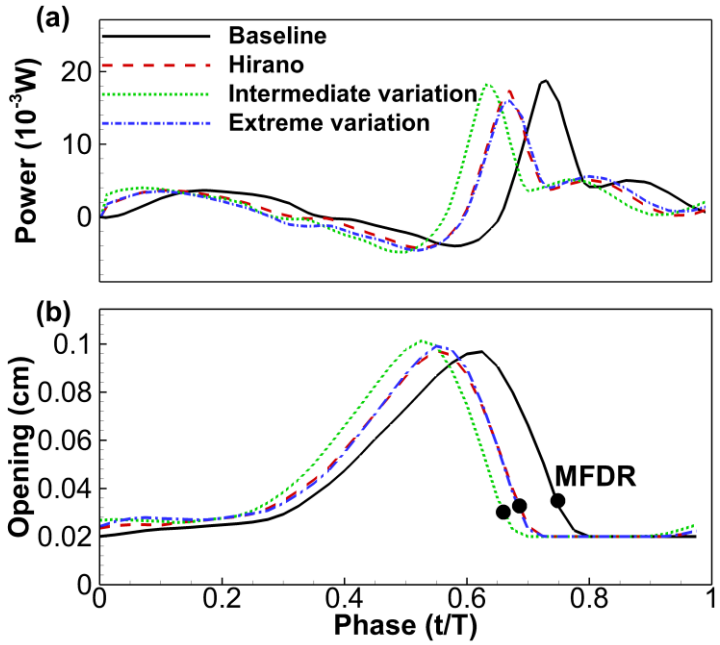


Figure 5.9 The phase-averaged energy transfer and glottal opening for the four cases. (a) The time history of the power transferred from the air flow to one vocal fold. (b) The glottal opening size. The time instants of MFDR are represented by black dots.

The cyclic net energy transfer to the vocal fold was calculated by integrating the power over one vibration cycle. It was 28.56±1.47μJ, 26.21±1.73μJ, 27.50±0.84μJ, 24.15±0.95μJ for the baseline, Hirano, intermediate variation, and extreme variation cases respectively. The second number denotes the root-mean-square variation between cycles. Hence the cyclic net energy transfer decreased by -8.2%, -3.7% and -15.4% in the Hirano model, intermediate variation model and extreme variation model, respectively, compared to the baseline model. The cyclic variation was 5.1%, 6.6%, 3.0% and 3.9% of the mean values, in the baseline model, Hirano model, intermediate variation model and extreme variation model, respectively. Figure 5.9 reflected that the effect on the cyclic net energy transfer might be associated with the positive peak of the power that occurred when the glottis was about to close. This peak contributed a significant amount of positive energy to the vocal fold. The changing trend of the peak among the cases was in accordance with the changing trend of the cyclic net energy transfer.

To explore the reason of the changes of the power peak among the cases, Figure 5.10 plots the phase-averaged flow pressure along the centerline of the glottis at the time instant when the positive power peak occurred for all the cases. It shows that, at this time instant, the flow pressure dropped to a negative value around -1 to -1.5kPa inside the glottis in all the cases. The baseline case had the lowest negative pressure, followed by the intermediate variation case, extreme variation case and the Hirano case in the ordered sequence. This low negative pressure generated a larger vacuum force to pull the vocal folds to close. Therefore, a lower negative pressure would contribute to higher power peak. Such sequence of intraglottal pressure is generally in accordance with the sequence of the power peak as shown in Figure 5.9 except that the Hirano case had the higher intraglottal pressure than the extreme variation case. It should be noticed that, other than the intraglottal pressure, the vocal fold velocity, face normal direction and the integration in time would also affect the energy consumption during vibration.

This might explain that though the Hirano case had the highest intraglottal pressure, the energy consumption was not the lowest. However, the trend that the longitudinal thickness variation would increase the intraglottal pressure which would further decrease the energy needed to sustain the vocal fold vibration was clear.

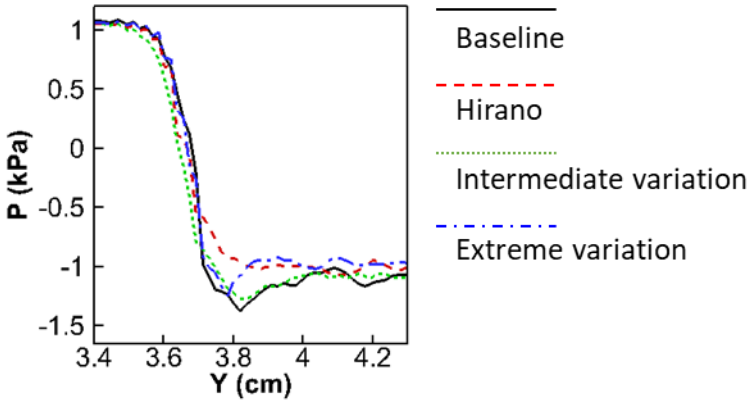


Figure 5.10 The phase-averaged intraglottal pressure along the centerline at the time instants of the positive power peak observed in Figure 5.9 for the four cases.

Then, it became interesting to explore the mechanism responsible for the lowest negative pressure in the baseline case. There are two primary mechanisms responsible for the negative intraglottal pressure during the closing phase. One is the steady effect that is related to the Bernoulli's effect, in which the pressure drop is proportional to the square of the flow rate. But such effect will be weak at the time instant when the flow rate was very low, such as the condition in the current cases. The second mechanism for the negative intraglottal pressure would be the unsteady effect which is related to the flow inertance. The pressure difference caused by the flow inertance in a tube can be calculated as $\Delta p = I(dQ/dt)$, where I is the flow inertance characterized by the density of the fluid, length and cross-section area of the tube, and Q is the volumetric flowrate. In contrast to the Bernoulli's effect

which is directly related to the flowrate, the flow inertance caused pressure change is related to the time derivative of the flow rate. During the closing phase of the glottis, the deceleration of the flow caused the negative time derivative of the flow rate and therefore the negative flow pressure inside the glottis. Figure 5.9 also denotes the time instants of the MFDR in the opening size plot for all the cases. It was found that the time instants of the positive power peak and MFDR were very close. Therefore, the very low intraglottal pressure observed in Figure 5.10 was primarily caused by the strong flow inertance effect due to the MFDR. Table 5-1 listed values of the MFDR for the cases. The baseline case had the highest MFDR value. Therefore, its intraglottal pressure was able to drop to a lowest value. That further explained the highest positive power peak in the baseline case.

5.4 Summary

The effect of longitudinal variation of vocal fold inner layer thickness on vocal fold vibration has been investigated using the three-dimensional fluid-structure interaction model. Four cases, one with the constant layer thickness in the longitudinal direction and three with different degrees of longitudinal variation, were studied and compared. In the extreme variation case, the thickest part of the cover/ligament layer can be 8.67/4.28 times the thinnest part. Such variation has caused 24% reduction in the stiffness at the middle of the vocal fold and 47% and 29% increase in the stiffness at the anterior and posterior ends, respectively. The main observations and discussions are summarized as below:

(1) The variation of inner layer thickness in the longitudinal direction as well as the resulted variation of stiffness distribution on the vocal fold did not affect the flow rate and vibration amplitude. The results implied that the flow rate and the vibration amplitude were highly related to the average stiffness over the vocal fold medial surface rather than the stiffness at the middle of the vocal fold. In our models, the stiffness at the middle of the vocal fold decreased up to 24%; however, the changes in

the average stiffness were only within 5%. The changes in the flow rate and vibration amplitude were within 6%.

(2) The maximum glottal angles were increased at the middle of the vocal fold and reduced at the anterior/posterior ends in the thickness varied cases. The results implied that the maximum convergent and divergent glottal angles along the longitudinal direction were largely determined by the local stiffness at the longitudinal positions. Such a relationship was more prominent in the divergent angles maybe because glottal closing is primarily driven by the elastic recoil force inside the vocal folds which are directly related to the stiffness.

(3) The thickness variations had little effect on the vocal fold vibration patterns. The proper orthogonal decomposition analysis showed nearly the same principle modes and energy distributions between the modes for all the models.

(4) The energy analysis revealed that a key contribution to the positive net energy transfer to the vocal fold was a positive power peak when the glottis was about to close. This positive power peak was primarily due to the negative intraglottal pressure caused by flow inertance which was related to MFDR. The thickness variations in the longitudinal direction were found to help decrease the energy needed in the sustained vibration due to the higher intraglottal pressures which were related with the smaller MFDR.

It is useful to point out that our results were based on the models with specific material properties. Vocal fold material properties vary from individual to individual depending on factors such as gender, race, age, and health. These variations are likely to affect the quantitative conclusions. For example, increasing the difference of Young's modulus between the cover and ligament layers will increase the variation of the stiffness along the longitudinal direction in the models, and this will affect

the quantities reported in the present study. There has been more experimental work recently to provide more comprehensive material measurements, including the anisotropic and nonhomogeneous aspects (Miri, 2014; Oren *et al.*, 2014; Dion *et al.*, 2016). These data can provide new insight in modeling vocal fold tissues.

CHAPTER 6 EFFECT OF COVER LAYER THICKNESS

In this chapter, the influence of vocal fold cover layer thickness on vocal fold vibratory dynamics has been studied. This work is reproduced from "Jiang, Weili, Xudong Zheng, and Qian Xue. 'Influence of vocal fold cover layer thickness on its vibratory dynamics during voice production.' *The Journal of the Acoustical Society of America* 146, no. 1 (2019): 369-380", with the permission of AIP Publishing.

6.1 Introduction

The cover layer thickness varies with vocal fold conditions. For instance, it increases in aging vocal folds of females (Hirano, Kurita and Sakaguchi, 1989). It is affected by voice warm-up and hydration.16 It can also be altered by vocal pathologies, such as polyps and nodules, and surgical treatments, such as dissection for treating cancers. Understanding how the cover layer thickness affects the vocal fold vibration and voice production will provide valuable insights into the diagnosis and treatment of related voice disorders.

Due to the difficulties of the measurement and control of vocal fold inner layers, past studies on the effect of vocal fold inner layer properties on voice production mostly relied on the theoretical, computational, and physical vocal fold models (Titze, Jiang and Drucker, 1988; Alipour *et al.*, 2011; Murray and Thomson, 2011; Xuan and Zhang, 2014; Zhang, 2017a). A number of studies have reported the effect of cover layer properties, including the thickness, damping, geometry, and elastic modulus, on the threshold pressure, a) onset frequency, and glottal closure(Titze, 1988; Zhang, 2009; Mendelsohn and Zhang, 2011). For the cover thickness, it is generally believed, based on the string formula, that a thicker cover layer results in a lower fundamental frequency(Titze, Jiang and Drucker, 1988; Titze, 2011). By using a theoretical surface wave model and a physical model of the vocal fold, Titze and his colleagues (Titze, Schmidt and Titze, 1995; Chan, Titze and Titze, 1997) found a positive dependence of

the phonation threshold pressure on the thickness of the cover layer. With a lumped mass model, (Gunter, 2004) showed that the mechanical stress during vocal fold collision was negatively correlated with the cover layer thickness. (Erath, Zañartu and Peterson, 2017) obtained the same conclusion by using a quasi, 1-D viscoelastic contact model. Using eigen analysis of a continuum vocal fold model, (Cook, Nauman and Mongeau, 2009) found that the eigenfrequencies were more influenced by layer elastic modulus rather than thickness. In another study on the eigenmodes of vocal fold, (Xue *et al.*, 2011) showed that the eigenmodes were not sensitive to the inner layer thickness either.

Despite these previous studies, the effect of the cover layer thickness on flow-induced vocal fold vibration has not been systematically studied. As the cover layer is closely associated with the propagation of the mucosal wave, we hypothesize that increasing the cover layer thickness would promote the formation and propagation of the mucosal wave, and so promote the excitation of the wavetype eigenmodes of the vocal fold. Such an effect would further affect the synchronization of the eigenmodes during fluid-structure interactions, and finally change the vocal fold vibration pattern, fundamental frequency, frequency spectrum, and voice outcomes. This study aimed to test this hypothesis through the current numerical approach. The acoustic solver was excluded as the focus is investigating the effect on the vocal fold dynamics. The cover-body thickness ratio was systematically varied in a wide physiologically-possible range. The effect of these variations on vocal fold stiffness, eigenfrequencies and modes, fundamental frequency, glottal flow rate, vocal fold vibratory dynamics, and the synchronization of the eigenmodes were analyzed in detail.

6.2 Simulation setup

Based on the model in Chapter 4, for the parametric study, seven different values of the cover thickness were considered, including T_c=0.5T_{co}, T_{co}, 1.5T_{co}, 2T_{co}, 2.5T_{co}, 3T_{co} and 5T_{co}. In these cases, the

ligament layer thickness was unchanged. The body layer thickness (T_b in Figure 6.1(a)) was adjusted accordingly to keep the total volume of the vocal fold unchanged. Hence these seven cases correspond to the cover-body thickness ratio (T_c/T_b) of 0.02, 0.05 (baseline), 0.07, 0.09, 0.12, 0.15 and 0.28, respectively. In the rest of the paper, the Tc/Tb value is used to represent the cases. Figure 6.1 (b) shows the inner layer profiles at the mid-coronal plane of the selective models. It should be noticed that T_c/T_b =0.28 ($5T_{co}$) might be an exaggerated value from a physiological point of view. However, the fact that a periodic doubling bifurcation occurred when T_c/T_b =0.15 ($3T_{co}$), which would be discussed later, indicated that the system was during a transition between two steady states and one of the steady stats occurred at T_c/T_b =0.28. That is the reason this case is included in the study. Although this case may be exaggerated physiologically, it provided a complete picture of the underlying mechanism driving the nonlinear dynamic transition of the system. Such knowledge would be very useful to understand the nonlinear dynamics observed in real vocal systems, such as those reported in (Berry *et al.*, 1996; Tokuda *et al.*, 2008; Zañartu *et al.*, 2011).

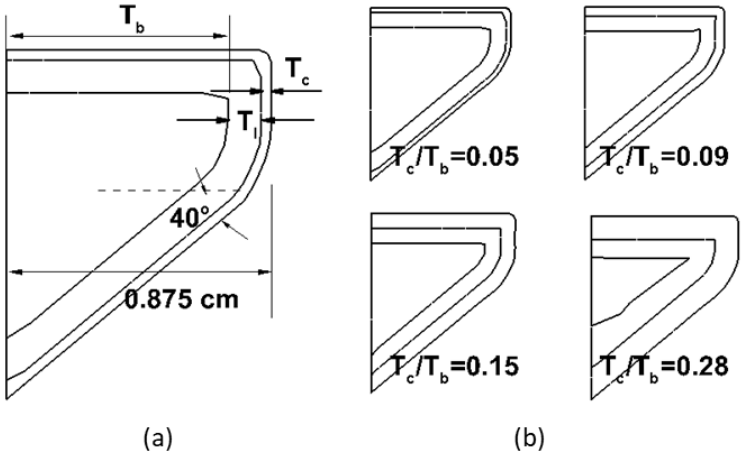


Figure 6.1 The cover thickness variation in the parametric cases. (a) A mid-coronal plane profile of the layered vocal fold showing the definition of the layer thicknesses (T_c, T_l, T_b) in the current study. They

add up to 0.875cm in all the cases. (b) The inner layer structures of the vocal fold in the selective cases in the parametric study.

6.3 Results and discussion

6.3.1 Effect on vocal fold stiffness

The cover layer has the lowest transverse Young's modulus among the three layers (Titze and Talkin, 1979; Kimura, Mau and Chan, 2011; Zhang and Luu, 2012). A thickened cover layer will result in a reduced stiffness of the vocal fold, and vice versa. It should be noted that, in this study, a distinction was made between the elastic property and stiffness. The elastic property is the material property just related with the material itself while the stiffness measures the local resistance under deformations. The stiffness is influenced by the elastic property as well as the geometric dimensions, loading direction and type of constraint. A numerical indentation technique was applied to measure the stiffness on the vocal fold surface. It was performed by applying a uniform pressure on a small area on the vocal fold surface and then numerically solving the static deformation using the finite element method. Let P be the pressure load, A be the loading area and Δx be the maximum deformation, the stiffness k was computed as $k=PA/\Delta x$. The details of this method are referred to (Geng, Xue and Zheng, 2017). In the current study, P=4kPa and A=1mm^2. The constant pressure loading was applied to create a loading condition comparable to that during fluid-structure interactions. The stiffness was measured at three locations at the mid-coronal plane of the vocal fold, including the superior and inferior side of the medial surface and a location on the subglottal surface, as shown in Figure 6.2 (a). Ideally, if the tissue surface is significantly larger than the size of the indentation (for example, a half infinite-surface), the stiffness only depends on the material properties, and if the material properties are uniform, it doesn't change along the surface. However, for the vocal fold model, its size is comparable to the indentation size,

therefore, its local stiffness is significantly affected by the geometries. The indentation depth varied with the location and the model. The smallest indentation depth was 0.26 mm and occurred at the inferior surface of the vocal fold model with the thinnest cover layer (T_c/T_b=0.02), while the largest indentation depth was 0.72 mm and occurred at the superior edge of the medial surface of the vocal fold model with the thickest cover layer (T_c/T_b=0.28).

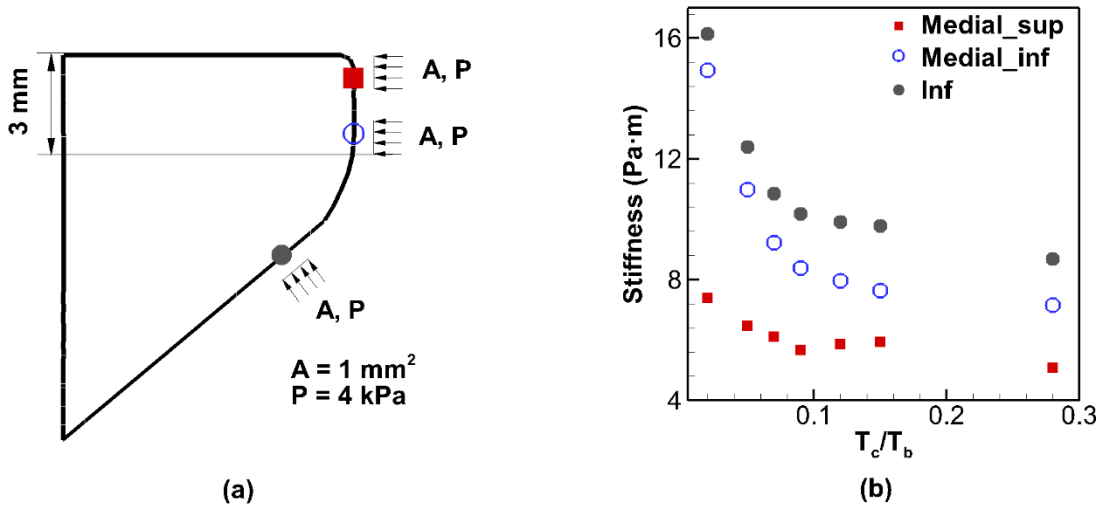


Figure 6.2 The stiffness measured at three locations at the mid-coronal plane of the vocal fold. (a) The three indentation locations in the superior and inferior sides of the medial surface and the inferior surface of the vocal fold. All of them were in the mid-coronal plane with the loading P=4kPa and loading area A=1mm^2. (b) The stiffness variation with cover-body layer thickness ratio in the three indentation locations.

Figure 6.2 (b) shows the stiffness values at the three locations and their variations with the cover-body layer thickness ratio. When the cover-body layer thickness ratio increased from 0.02 to 0.28, the stiffness decreased 31.2% at the superior side of the medial surface, 52.1% at the inferior side of the

medial surface, and 46.1% at the inferior surface. The decreasing rate of the stiffness decreased with the increasing cover-body thickness ratio, indicating a reduced sensitivity of the stiffness to an increased cover-body thickness ratio. It was also observed that the stiffness increased from the superior to the inferior. It was caused by two factors. The first is related to the asymmetric geometry of the vocal fold from inferior to superior and that the extra subglottal part provided more resistance to the loading according to (Geng, Xue and Zheng, 2017). The second is related to the fact that some parts of the superior measurement area only had the resistance from the cover layer below it while all the inferior measurement area had the resistance from the cover, ligament and body layers.

6.3.2 Effect on eigenmodes and eigenfrequencies

Since the ultimate vibration pattern of the vocal fold is determined by the shape of the excited eigenmodes, especially the low order modes, the effect of the cover-body layer thickness ratio on the eigenmodes and eigenfrequencies was evaluated. Figure 6.3 (a) displays the profile of the mid-coronal plane of the vocal fold at the two extreme positions of the five lowest-eigenfrequency modes of the baseline model. The dashed line represents the original shape and the solid lines represent the deformed shapes. The modes are ordered in a way that the eigenfrequency increases with the mode number. Mode 1 and 2 primarily represented a vertical motion and a medial-lateral motion of the vocal fold, respectively. Starting from Mode 3, the wave motions were observed on the vocal fold surface. From Mode 3 to Mode 5, the wavelength gradually decreased with the mode order. Moreover, in Mode 3, the wave was primarily on the superior surface of the vocal fold, while, in Mode 4 and 5, the waves were on the entire vocal fold surface. A close look at the wave patterns in Mode 4 and Mode 5 revealed that these two modes generated an out-of-phase motion between the superior and inferior edges of the

medial surface of the vocal fold, which was essential for the energy transfer from the airflow to the vocal fold.

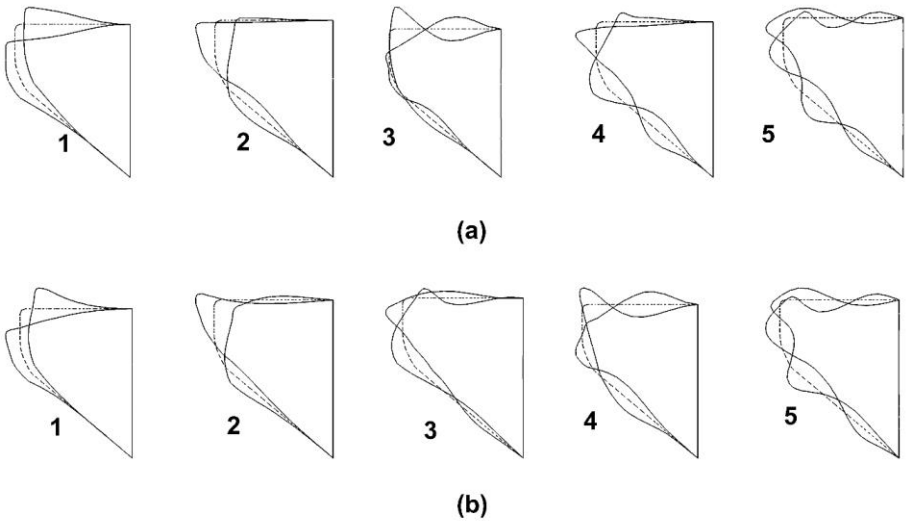


Figure 6.3 Profiles of the mid-coronal plane of the vocal fold for eigenmodes 1-5. (a) the baseline case; (b) Tc/Tb=0.28 case. The solid lines represent the two extreme deformed shapes and the dashed line represents the original shape.

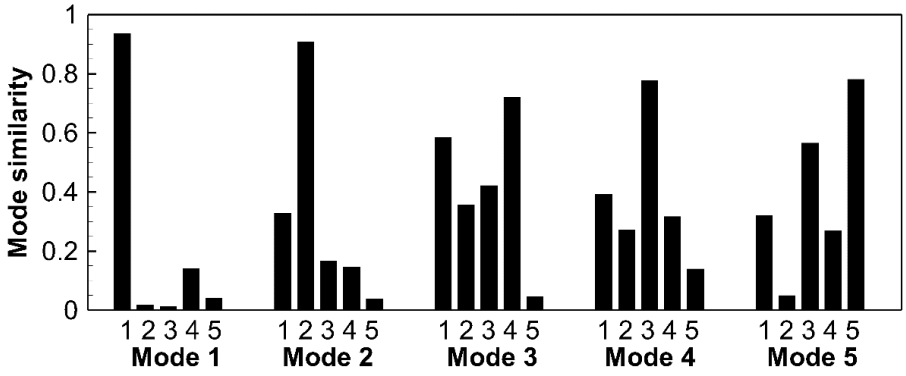


Figure 6.4 The similarity of the eigenmodes 1-5 between the T_c/T_b =0.28 case and baseline case. The *i*th bar in the category of Mode j represents the dot product between the *j*th mode in the baseline case and the ith mode in the T_c/T_b =0.28 case.

For the cases where the cover-body thickness ratio was between 0.02 and 0.15, the eigenmodes were very similar to those in the baseline case. In the case T_c/T_b =0.28, the eigenmodes were different. Figure 6.3 (b) shows the shape of the five lowest-eigenfrequency modes for the T_c/T_b =0.28 case. A close comparison of these modes to those in the baseline case revealed that Mode 1, 2, 5 were highly similar with the corresponding modes in the baseline case, while Mode 3 and Mode 4 switched the modal order. Specifically, Mode 3 and 4 in this case were highly similar with Mode 4 and 3 in the baseline case, respectively. To precisely quantify the similarity between the modes, the dot-product between the eigenmodes of the two cases were computed (Xue *et al.*, 2011). The method is described below. The j^{th} mode was written into a single vector $\overrightarrow{q_j} = (u_{j1}, v_{j1}, w_{j1} \dots u_{jN}, v_{jN}, w_{jN})$, where u_{jm}, v_{jm} and w_{jm} represent the displacements of the m^{th} grid point in the *x*, *y* and *z* directions, respectively. This vector is normalized with its own magnitude leading to a normalized vector: $\overrightarrow{Q_j} = \overrightarrow{q_j}/|\overrightarrow{q_j}|$, where $\left|\overrightarrow{q_j}\right| = \sqrt{\sum_{i=1}^{N}(u_{ji}^2 + v_{ji}^2 + w_{ji}^2)}$ and N is the total number of grid points. The dot-product of any two normalized modes is indicative of the similarity between the two modes with the value of one corresponding to an exact match, and zero indicating orthogonality. Figure 6.4 shows the dot-products between each mode in the T_c/T_b =0.28 case and each mode in the baseline case. It is clear that Mode 1, 2 and 5 in the T_c/T_b =0.28 case were highly similar to the corresponding modes in the baseline case with the dot-products ranging between 0.79 and 0.94. Mode 3 and 4 in the T_c/T_b =0.28 case were highly similar to Mode 4 and 3 in the baseline case with the dot-product value of 0.72 and 0.78, respectively. Since Mode 4 was the lowest-eigenfrequency mode in the baseline case presenting the elastic wave on the entire vocal fold surface, its moving-up in the modal order in the T_c/T_b =0.28 case implied that a very thick cover would promote the formation of the elastic waves on the vocal fold surface.

Figure 6.5 shows the value of eigenfrequencies of mode 1-5 in each case and their variations with the change of the cover-body thickness ratio. The eigenfrequencies decreased with the increasing cover-body thickness ratio. This was due to the reduced vocal fold stiffness.

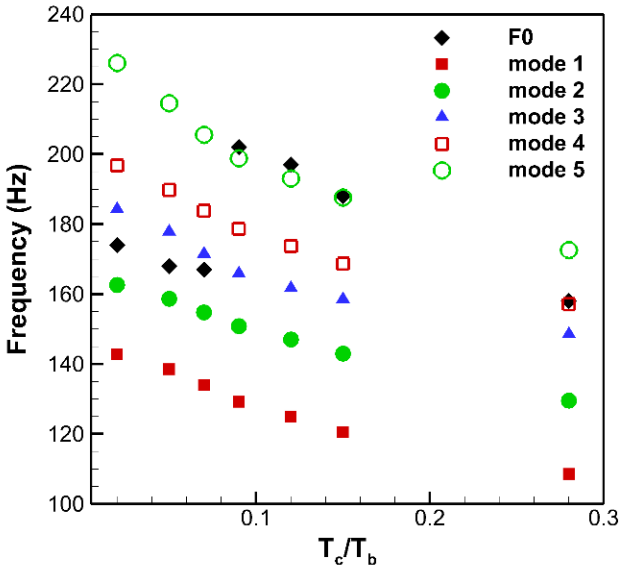


Figure 6.5 The eigenfrequencies of mode 1-5 and the fundamental frequency (F_0) in each case.

6.3.3 Effect on glottal flow rate and spectrum

Figure 6.6 shows the time variation of the volumetric flow rate through the glottis in each case. Interestingly, it was seen that when the cover-body thickness ratio was low (T_c/T_b =0.02-0.07), the flow rate reached the periodic, steady cycles very quickly. When the cover-body thickness ratio was increased to T_c/T_b =0.09, the waveform of the flow rate became irregular and non-periodic with large variations in the frequencies and peak values from cycle to cycle. The irregularity continued in the T_c/T_b =0.12 case. When the thickness ratio was increased to T_c/T_b =0.15, the flow rate showed a periodic waveform again, but with alternating low and high peaks, indicating a period-doubling behavior characterized by two alternating peaks in each cycle. Keeping increasing the thickness ratio to T_c/T_b

=0.28, the system was back to the regular and periodic vibration again, but with a much longer transition time to reach the steady cycles compared to the baseline case.

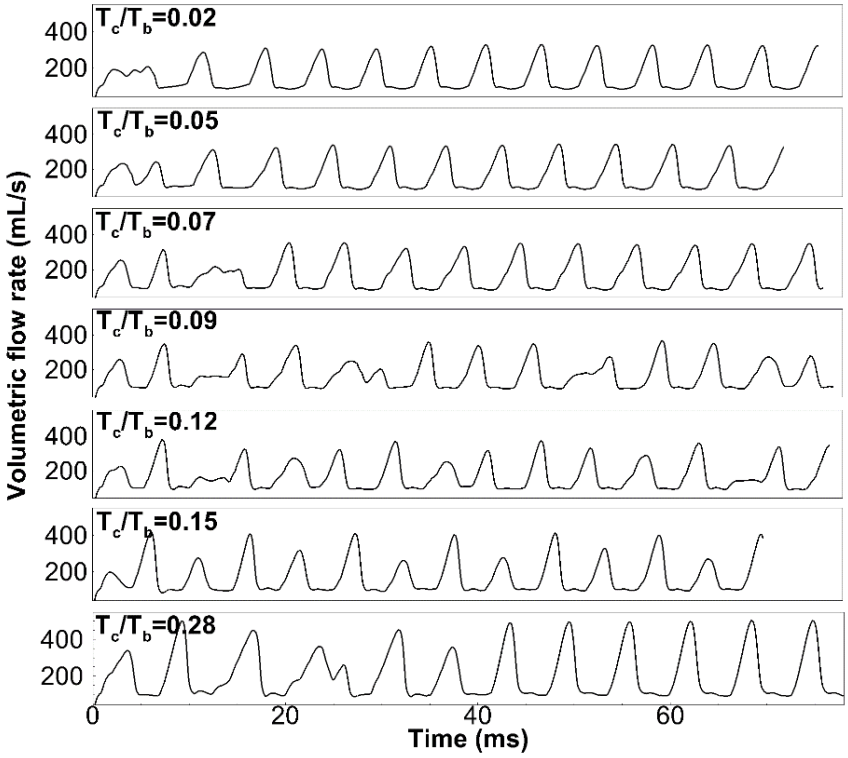


Figure 6.6 The time history of the volumetric glottal flow rate in each case.

The changes of the flow waveforms clearly indicated a nonlinear bifurcation of the system associated with the changes of the cover-body thickness ratio. Period-doubling is often associated with the appearance of a subharmonic in the frequency spectrum. To examine this behavior, Figure 6.7 shows the spectrum of the flow rate in each case. The spectrum was obtained through performing fast Fourier transform in all the cases with the time period of about 0.05 s, indicating a frequency resolution of about 19Hz. The values of the fundamental frequency and signal-to-noise ratio (SNR) are also shown in each sub-figure. The spectrum and SNR were calculated based on the data of the steady cycles or late cycles if there was no steady state. In the cases where the cover-body thickness ratio was between 0.02

and 0.07, the spectrum showed a typical pattern of limit-cycle oscillations, dominated by the fundamental frequency and its superharmonics. The SNR deceased with the increasing cover-body thickness ratio. In the T_c/T_b =0.09 case, the spectrum became noisy. The SNR decreased by about 53% from the baseline case. The spectrum showed two nearly equal peaks at 165Hz and 201Hz with the peak at 201 Hz slightly higher than the peak at 165 Hz. The superharmonics of the peaks could no longer be distinguished. This case indicated the onset of the bifurcation. In the T_c/T_b =0.12 case, the spectrum was still noisy. The SNR further decreased to just about 12% of that in the baseline case. In the T_c/T_b =0.15 case, the spectrum clearly showed a 1/2 subharmonic frequency (frequency below the fundamental frequency in a ratio of 1/2) and it was dominated by the fundamental frequency, the subharmonic and the linear combination of the two. This case indicated the complete of the period-doubling bifurcation. The system was stable again at this new state and the SNR increased to 42% of that in the baseline case. Further increasing the cover-body thickness ratio to 0.28 created another bifurcation that the system returned to the periodic, limit-cycle oscillation with the spectrum dominated by the fundamental frequency and its superharmonics again and the SNR as high as 96% of that in the baseline case.

6.3.4 Effect on the fundamental frequency and mode synchronization

The fundamental frequency of each case is superimposed in Figure 6.5 using the diamond symbols. It was interesting that, while the fundamental frequency decreased when T_c/T_b was increased from 0.05 to 0.07 and from 0.09 to 0.28, which was as expected since the stiffness of the vocal fold was reduced, it increased when the thickness ratio was increased from 0.07 to 0.09. This singular behavior violates the general understanding that the fundamental frequency decreases with the decreasing stiffness of the vocal fold based on the string formula. By noticing that the case T_c/T_b =0.09 was the onset of the period-doubling bifurcation as observed in Figure 6.7, we hypothesized that the period-

doubling bifurcation and the jump of the fundamental frequency were related and their occurrence were due to the increased strength of the higher-frequency modes in the modal synchronization process that broke the initial modal synchronization and attracted the modes to synchronize at higher frequencies.

To test the hypothesis, we computed the normalized modal coefficient of each eigenmode in each case. The coefficient reflected the contribution of each mode to the unit length of the displacement vector of the vocal fold vibration. The approach is described below. The vocal fold vibration was assumed to be a linear superposition of the eigenmodes as shown in (6.1

$$V(t) = c_1(t)V_1 + c_2(t)V_2 + \cdots + c_n(t)V_n \tag{6.1}$$

where V(t) is the vocal fold displacement vector at the time instant of t, V_i (i=1~n) is the ith eigenmode and c_i (t) is the corresponding eigen coefficient that is a function of time as it is different at different time steps. Based on this assumption, the vocal fold displacement vector was transferred from the general three-dimensional physical coordinates system to a modal coordinates system. The modal coefficients were the amplitudes of each modal component. It needs to be pointed out the equilibrium position of the vocal fold during vibration was shifted from the initial rest position due to the subglottal pressure. Therefore, the displacement vector, V, was computed based on the equilibrium position which was the average of the coordinates of the vocal fold nodes throughout the steady cycles. Based on the orthogonality of the eigenmodes ($V_j[M]V_i = \delta_{ij}$), the modal coefficients were computed as

$$c_i(t) = V(t)[M]V_i \tag{6.2}$$

where [M] is the global mass matrix. The modal coefficients were further normalized by the magnitude of the displacement vector as $c_i'(t) = c_i(t)/\sqrt{|V(t) \cdot V(t)|}$. The modal coefficients represented the amplitudes of the vibration of each mode per unit length of the displacement vector.

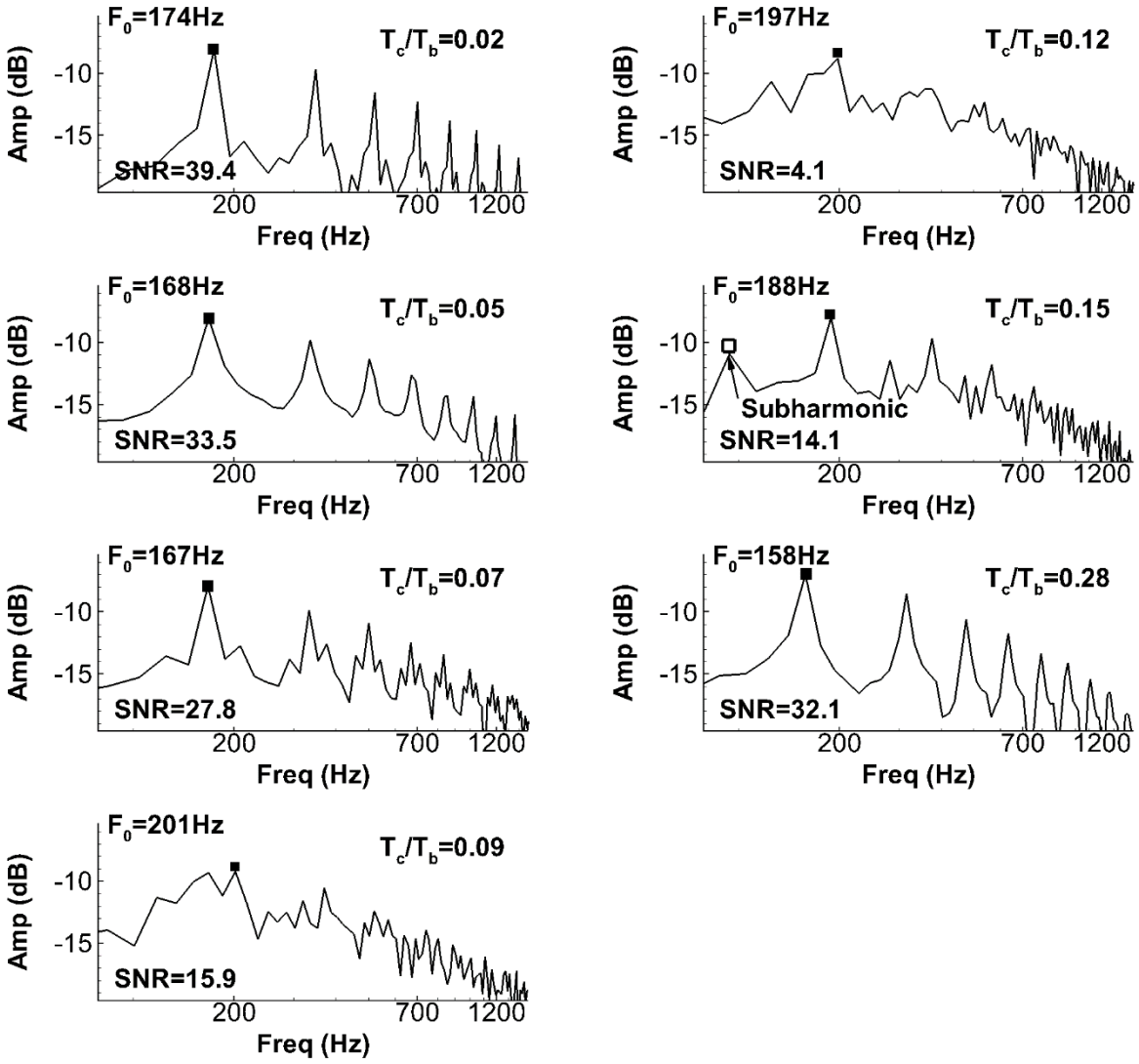


Figure 6.7 The spectrum of the flow rate waveform in each case. The fundamental frequency (F_0) and signal-to-noise ratio (SNR) of each case are denoted in the subfigures. The emergence of the subharmonic is highlighted in the T_c/T_b =0.15 case.

Figure 6.8 shows the root-mean-square (RMS) values of $c_i'(t)$ of the five lowest-eigenfrequency modes in each case. It was found that when T_c/T_b was between 0.02 and 0.07, the amplitudes of Mode 1, 2, 3 were much higher than those of Mode 4, 5. According to Figure 6.5, the fundamental frequencies

of these cases were close to the eigen frequencies of Mode 2 and 3. It was further found that in these cases, the amplitudes of Mode 2 and 4 decreased with the increase of the cover-body thickness ratio while that of Mode 3 increased. However, when the thickness ratio was increased from 0.09 to 0.15, the amplitudes of Mode 4 and 5 quickly increased with the amplitude of Mode 3 dropping quickly and the amplitudes of Mode 1 and 2 remaining nearly unchanged. This behavior confirmed our hypothesis that the jump of the fundamental frequency and the chaotic vibration at T_c/T_b =0.09 were associated with the increase of the strength of the higher-frequency modes. With the continuous increase of the cover-body thickness ratio, the amplitudes of Mode 4 and 5 continued to increase and the amplitude of Mode3 continued to drop. At T_c/T_b =0.15, they had nearly the same amplitudes, indicating a comparable strength among them, which may be the reason for the periodic doubling vibration which has two energy peaks (fundamental and its subharmonics). With the further increase of the cover-body thickness ratio, at T_c/T_b =0.28, although the amplitudes of Mode 4 only increased slightly, the amplitude of Mode 3 dropped significantly; therefore, the relative strength of Mode 4 and 5 was still increased. This has led the system back to a new regular periodic vibration state but with the fundamental frequency near the eigenfrequency of the higher-frequency modes. As Mode 2 represents a medial-lateral motion involving all the layers and Mode 4 and 5 represent wave motions on the vocal fold surface, the energy decrease of Mode 2 when T_c/T_b increased from 0.02 to 0.07 and the energy increase of Mode 4 and 5 when T_c/T_b increased from 0.07 to 0.28 suggest that increasing the cover-body thickness ratio would make the vibration more concentrated on the superficial layer of the vocal fold and generate higher frequencies. It may have implications to modal-falsetto transition as high-frequency falsetto is mostly with cover and ligament layer vibrating and low-frequency modal register also involves the body layer. The results imply

that the vocal fold vibration in falsetto might be dominated by the higher-order wave-type of eigen modes and increasing the cover-body thickness ratio would be easier to generate falsetto.

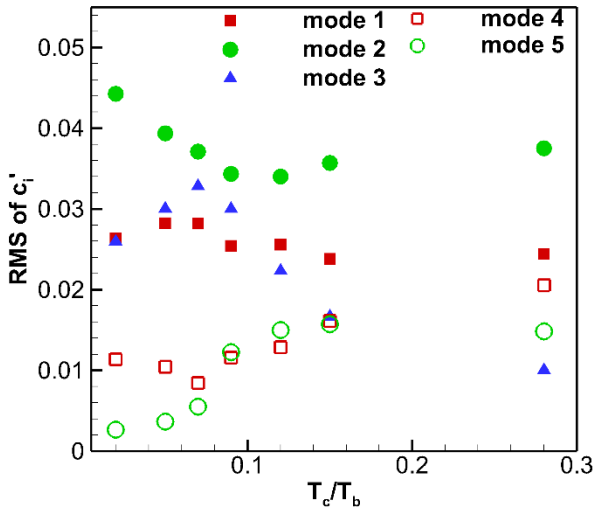


Figure 6.8 The root-mean-square values of c_i' of eigenmodes 1-5 in each case. c_i' stands for the amplitude of the vibration of each mode per unit length of the displacement vector.

Figure 6.9 shows the frequency spectrum of the modal coefficients of the five lowest-eigenfrequency modes in each case. The frequency resolution of all the cases is about 19Hz which is the same with the ones in Figure 6.7. It will allow us to examine the vibration frequency of each eigenmode so that the modal synchronization can be analyzed. The five lowest eigenfrequencies were also indicated by the vertical lines in each subfigure to show the relationship between the vibration frequencies and the eigenfrequencies. The spectra clearly show that, for the cases where the cover-body thickness ratio was between 0.02 and 0.07, all the modes were vibrating at the same frequencies. When T_c/T_b was increased to 0.09, the modes started to show different spectrum patterns, indicating that the modal

synchronization was broken. It was noticed that, for all the modes, there were two dominant peaks, one close to the eigenfrequency of Mode 3 and the other close to the eigenfrequency of Mode 5, indicating

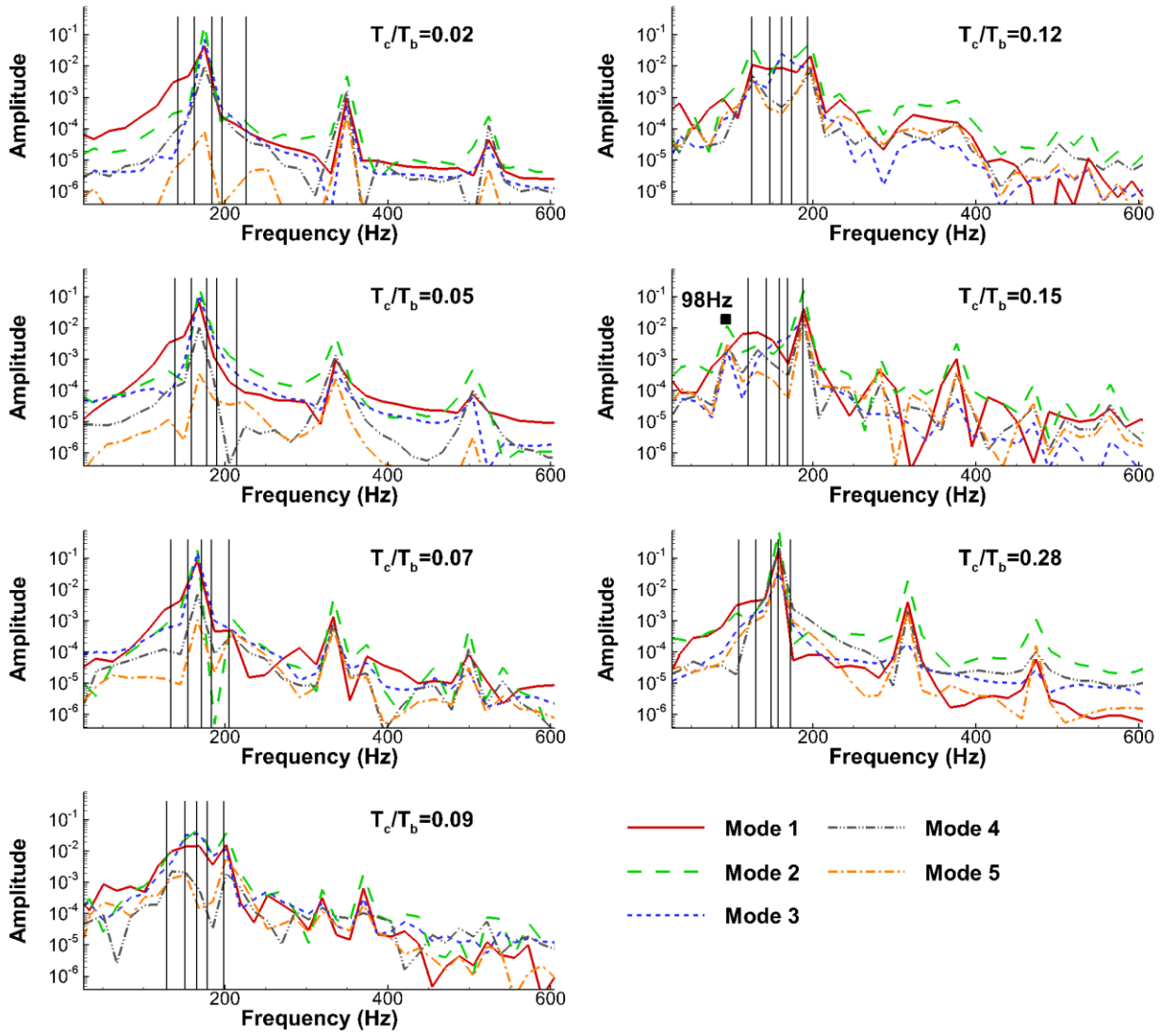


Figure 6.9 The spectrum of each modal coefficient (c_i) of eigenmodes 1-5 in each case. The vertical lines in each subfigure indicate the eigenfrequencies of eigenmodes 1-5 in the corresponding case. The frequency of the subharmonic (98Hz) is denoted in the T_c/T_b =0.15 case.

that the higher-frequency modes (Mode 4 and 5) started to act as another basin of attraction for the system. This was due to the increased strength of the two modes, as seen in Figure 6.8. At T_c/T_b =0.12, the system showed the similar two-peak behavior except that the two peaks were close to the eigenfrequency of Mode 1 and 5, respectively. When T_c/T_b =0.15, the two-peak behavior vanished and all the modes only showed one peak which was close to the eigenfrequency of Mode 5. Moreover, Mode 2-5 all have energy concentration at the subharmonic frequency (98Hz). With the further increase of the energy portion of Mode 4 and 5, at T_c/T_b =0.28, all the modes were entrained to the same frequency close to the eigenfrequency of Mode 4.

In summary, the increased cover-body thickness ratio has promoted the excitation of the wave-type eigenmodes of the vocal fold (Mode 4 and 5), which gradually became a stronger basin of attraction for the system, driving the system to a new periodic doubling vibration state and then back to the regular periodic vibration with higher fundamental frequencies. During the transition between the two steady states, the modes were vibrating at different vibration frequencies, resulting in complex chaotic and period-doubling vibrations.

6.3.5 Effect on vocal fold vibration pattern

Figure 6.10 shows the phase-plane plot of the lateral displacements of the points at the superior and inferior sides of the medial surface in the mid-coronal plane of the vocal fold in each case. The positions of the two points are denoted in Figure 6.2(a). It clearly shows that when T_c/T_b was between 0.02 and 0.07, the systems underwent the limit-cycle vibrations. At T_c/T_b =0.09 and 0.12, the systems underwent the chaotic vibrations with both the vibration amplitude and pattern varying from cycle to cycle. At T_c/T_b =0.15, the system showed the period-doubling vibration with the alternative larger and smaller vibration amplitudes. At T_c/T_b =0.28, the limit-cycle vibration resumed. Throughout all the cases,

the vibration amplitude at both points increased with the increasing cover-body thickness ratio in accordance with the decreased stiffness on the vocal fold surface. Moreover, a 45° dash line was superimposed in the figures to examine the glottal shape change. The points on the 45° line indicate a straight glottal channel, the points below the line indicate a divergent channel and the points above the line indicate a convergent channel. Moreover, the distance from the point to the 45° line reflected the glottal angle. When the point was more away from the line, the glottal angle was larger. Figure 6.10 reflected that, with the increasing cover-body thickness ratio, the phase plot gradually shifted from upper left to lower right. This change indicated a gradually decreased convergent angle and increased divergent angle during vocal fold vibration. By carefully examining the vibrations of the vocal folds, it was noticed that this effect was due to the larger opening of the superior part of the vocal folds with the

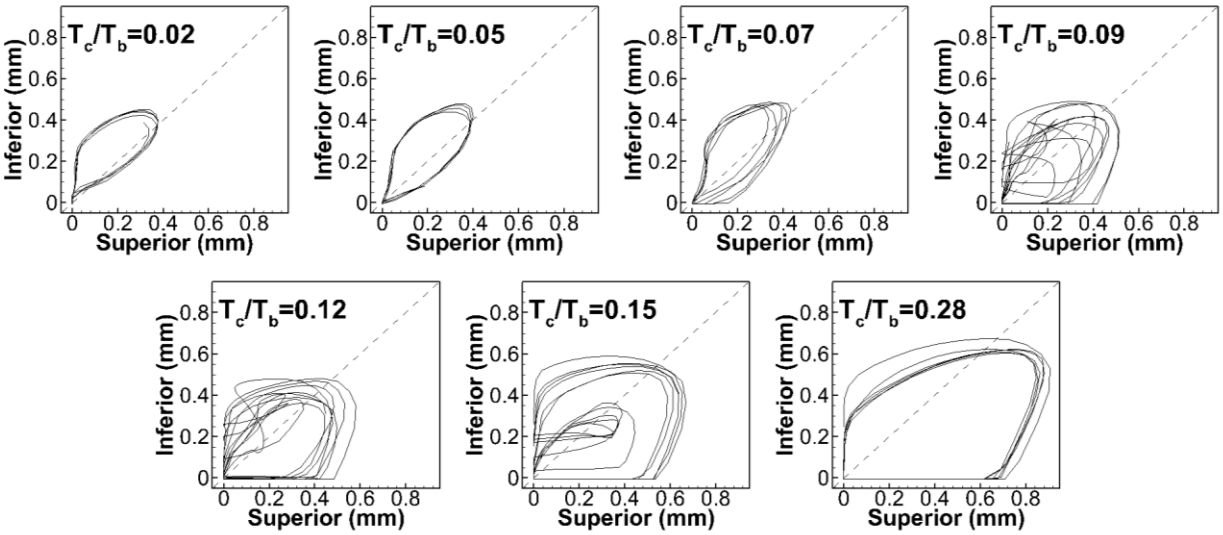


Figure 6.10 The phase-plane plot of the lateral displacement of the points at the superior and inferior sides of the mid-coronal medial surface of the vocal fold in each case. The positions of the two points are denoted in Figure 6.2 (a). The 45° dashed line indicates a straight glottal channel shape. The plot goes clockwise direction chronologically.

increased cover-body thickness ratio. Briefly, when the cover-body thickness ratio was low (between 0.02 and 0.07), the inferior and superior parts of the vocal fold reached their maximum openings at almost the same time when the glottis was straight. After that, they both started to close. When the cover-body thickness ratio was increased, when the inferior part had reached the maximum and was about to close, the superior was still in the opening motion. This has resulted in a longer time of a divergent glottis as well as a larger maximum divergent angle. Therefore, the phase plot was shifted to the right.

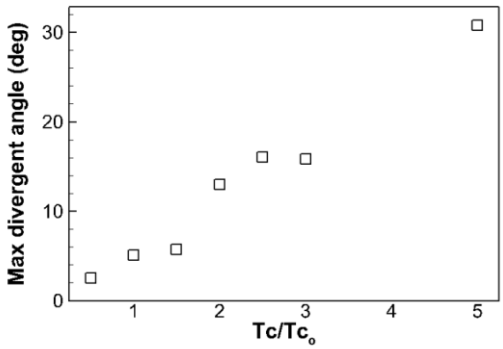


Figure 6.11 The maximum divergent angle during vocal fold vibration of each case.

Figure 6.11 plots the maximum divergent angle between the two vocal folds in each case. The angle was calculated by tracing the lateral displacement of the inferior and superior side points. The plot clearly shows that the maximum divergent angle increased with the increasing cover-body thickness ratio. When the cover-body thickness ratio was between 0.02 and 0.07, the angle was below or around 8 degrees. At T_c/T_b =0.09, the angle was quickly increased to 13 degrees. When T_c/T_b =0.12 and 0.15, it was further increased to 16 degrees. At T_c/T_b =0.28, the angle had another quick increase to a value as high as 31 degrees. The two quick increases of the maximum divergent angle occurred at the onset of the two bifurcations. As the larger divergent angle indicated larger amplitude of the mucosal wave

propagation, such effect was in accordance with our previous analysis that the increase of the cover-body thickness ratio excited the wave-type eigenmodes, which had caused the bifurcations of the dynamics of vocal fold vibrations.

6.4 Summary

The cover-body layer thickness ratio was systematically varied in a wide range. The effect of the cover-body layer thickness ratio on the vocal fold stiffness, eigenfrequencies and eigenmodes, fundamental frequencies, glottal flow rate, vocal fold vibratory dynamics and synchronization of the eigenmodes were analyzed in detail. The main conclusions are summarized below.

As the cover-body thickness ratio increased, the stiffness of the vocal fold decreased. This resulted in the decrease of the eigenfrequencies and the increase of the vibration amplitude of the vocal fold. A significant increase of the cover-body thickness ratio (T_c/T_b =0.28 in the present study) also affected the order of the eigenmodes in which the wave-type eigenmodes moved up with lower eigenfrequencies.

The cover-body layer thickness ratio significantly affected the strength and synchronization of the eigenmodes during vibration, and ultimately affected the fundamental frequency and vocal fold vibration pattern. It was found that when the cover-body thickness ratio was low, the vocal fold vibration was dominated by the eigenmodes 1-3, which synchronized all the eigenmodes to a fundamental frequency near their eigenfrequencies. Because these three modes were not wave-type modes in the medial surface, they did not create a strong mucosal wave on the vocal fold surface; therefore, the divergent angle of the vocal fold vibration was small. In these cases, the fundamental frequency decreased with the increasing cover-body thickness ratio. When the cover-body thickness ratio was increased beyond a critical value (0.09 in this work), the strength of the wave-type higher-

eigenfrequency modes was significantly increased. These modes became a stronger basin of attraction, breaking the initial synchronization of the modes and attracting other modes to vibrate at higher frequencies. As a result, the modes started to vibrate at different frequencies, resulting in chaotic vibrations. Also, because more energy was shifted to higher frequencies, the fundamental frequency was increased. Further increasing the cover-body thickness ratio continued to increase the strength of the higher-frequency modes and created a new steady state where the system underwent a period-doubling vibration. At this state, all the modes were synchronized at the fundamental frequency and its ½ subharmonics. The presence of the subharmonic was likely due to the comparable strength of eigenmodes 1-3 and higher-eigenfrequency modes. At last, further increasing the cover-body thickness ratio continuously increased the strength of the higher-eigenfrequency modes, resulting in a new steady regular periodic vibration where all the modes were synchronized at the fundamental frequency near the higher eigenfrequencies. During this process, because of the continuous increase of the strength of the wave-type modes, the mucosal wave on the vocal fold surface became stronger, and so the maximum divergent angle of the glottis was increased.

In summary, our simulation results showed that the cover-body thickness ratio had a complex nonlinear effect on the vocal fold vibration and voice production. Increasing the cover-body thickness ratio would promote the excitation of the wave-type modes of the vocal fold, which were also higher-eigenfrequency modes. These modes became a stronger basin of attraction for the system, driving all the modes to vibrate at higher frequencies, creating a period-doubling bifurcation state and eventually a new steady regular vibration at a fundamental frequency near the higher eigenfrequencies. Because of the increased strength of the higher-eigenfrequency modes, the fundamental frequency could have a sudden increase with the increasing cover-body thickness ratio although the stiffness of the vocal fold

was decreased. During the transition, the modes vibrated at different vibration frequencies, resulting in chaotic vibrations. The increased cover-body thickness ratio also resulted in the increase of the maximum divergent angles during the vibration.

In the end, it is useful to point out the limitations of the study. First, the study was based on one set of model parameters including the geometry of the vocal fold and inner structures, the lung pressure and the elastic properties. Different choices of these parameters may result in different system behaviors. For example, a higher or lower value of the lung pressure might result in the transition happening at different values of the cover-body thickness ratio. Moreover, changing the Young's moduli might change the eigenmode shape and order, onset of the nonlinear bifurcation, bifurcation type, vocal fold vibration pattern, etc. However, the main finding of the study was that a larger cover-body layer thickness ratio would reduce the stiffness on the vocal fold surface and promote the excitation of the wave-type eigenmodes during vocal fold vibration, which could cause complex nonlinear behaviors. This effect was due to that the elastic modulus of the cover layer is smaller than that of the other two layers. We think the conclusion has general validity in normal phonation condition because the elastic modulus of the cover layer is normally lower than the other two layers. Second, the effect of intrinsic laryngeal muscle activation was not considered due to the complexity of the model. The stiffness of the vocal fold would be significantly affected by the muscle activations and it might change the eigenmode shape and order, onset of the nonlinear bifurcation, bifurcation type, vocal fold vibration pattern, etc. Lastly, the vocal fold model was based on the three-layer assumption. Ideally, a more complicated vocal fold model which consists of the five histological layers with their accurate mechanical properties would provide more physiologically realistic predictions on vocal fold vibration patterns. For the cover layer, it was reported that the epithelium has a much larger elastic modulus than the superficial layer of the lamina

propria(Hirano and Kakita, 1985). The experiments on synthetic vocal folds(Murray and Thomson, 2011; Xuan and Zhang, 2014) found that adding a stiff epithelium layer with an elastic modulus higher than the other layers would improve the closure and mucosal wave of vocal fold vibrations. However, there exists very little measurement of the mechanical properties of vocal fold tissues based on the five layers. Most of the measurement were based on either the two-layer or three-layer assumption(Miri, 2014; Dion *et al.*, 2016). Furthermore, the focus of the current study was on the effect of cover layer thickness whose variation is mostly due to the changes in the superficial layer of the lamina propria. Based on these considerations, we think the three-layer assumption is reasonable for the current study. However, it is important to study the effect of the epithelium in the future.

CHAPTER 7 CONCLUSION

The voice production is a complex process. The current study focuses mainly on the generation of the sound source- the fluid-structure interaction process in the larynx. A high-fidelity in-house numerical solver, including a sharp-interface immersed boundary method based incompressible fluid solver, a hydrodynamic/acoustics splitting method based acoustics solver and a viscoelastic solid solver has been employed to simulate the flow-induced vocal fold vibrations. The numerical approach has been first validated against experiment measurements and then employed in the study of the effect of the longitudinal variation of layer thickness and the effect of the cover-body thickness ratio. Main observations and discussions are summarized as below:

(1) Subject-specific modellings with canine larynx and pigeon syringes have been provided as validation of the numerical approach. Specifically, the pigeon study is a strict one-to-one comparison that the tissue geometry, material property and tissue dynamics all come from one single subject and a blinding procedure has been employed that the geometry or material properties are not adjusted according to dynamic performance. Glottal flow waveform, vocal apparatus dynamics (vocal fold divergent angles, vibration shape, mucosal wave speed), acoustic features (fundamental frequency, source level, spectral slope and acoustic power) and intraglottal velocity fields (glottal jet dynamics, swirl strength) have been compared between the simulation results and the experiment measurement and showed good agreement. This validation demonstrated that the model could reproduce realistic phonatory dynamics during voice production.

(2) The numerical solver has been applied in a simplified human vocal fold and vocal tract system. With fluid-structure-acoustic interaction simulation, the result shows that a reasonable glottal flow waveform were captured by the model and important voice-quality associated parameters were

found to be well within the normal physiological ranges. The important convergent-divergent vibration pattern of vocal folds was captured. POD analysis demonstrated the 1:1 entrainment of the two dominant vibratory modes. The analogy between the vocal tract and a quarter-wave resonator was demonstrated. The simulation result suggested a strong source-filter interaction since the acoustic perturbed flux and pressure inside the glottis as well as the supraglottal tract were all at the same order with their incompressible counterparts.

(3) Based on the model in (2), the effect of longitudinal variation of cover and ligament layers thicknesses on vibration has been investigated. Four cases with thickness ratio ranges from 1 to 8.67 for the cover layer and 1 to 4.28 for the ligament layer have been calculated. The results show that the variation of inner layer thickness in the longitudinal direction as well as the resulted variation of stiffness distribution on the vocal fold have little effect on the flow rate, vocal fold vibration amplitude and pattern (POD modes). The small effect on flow rate and the vibration amplitude might be due to the relatively small change in the average stiffness. However, glottal angle, which is more affected by the local stiffness, has been increased around the midcoronal plane but decreased near two ends in high thickness ratio cases. Energy analysis shows that the longitudinal thickness variations helped decrease the energy needed in the sustained vibration due to the higher intraglottal pressures during closing.

(4) Still based on the model in (2), the cover-body thickness ratio has been increased from 0.02 to 0.28, corresponding to the cover layer thickness changes from 0.5 to 5 times of the baseline value. The results showed that the cover-body thickness ratio had a complex nonlinear effect on the vocal fold vibration and voice production. Increasing the cover-body thickness ratio would promote the excitation of the wave-type modes of the vocal fold, which were also higher-eigenfrequency modes. These modes became a stronger basin of attraction for the system, driving all the modes to vibrate at higher

frequencies, creating a period-doubling bifurcation state and eventually a new steady regular vibration at a fundamental frequency near the higher eigenfrequencies. Because of the increased strength of the higher-eigenfrequency modes, the fundamental frequency could have a sudden increase with the increasing cover-body thickness ratio although the average stiffness of the vocal fold was decreased. During the transition, the modes vibrated at different vibration frequencies, resulting in chaotic vibrations. The increased cover-body thickness ratio also resulted in the increase of the maximum divergent angles during the vibration.

The limitations regarding each specific study have been discussed in the end of each chapter. A summary is provided here:

(1) Due to the model complexity and the high numerical cost, the parametric study was still in a relatively small range. For example, in Chapter 6, though the cover layer thickness already changed from 0.5 to 5 times of the baseline value, the material property remained the same in the whole process. The result would be more general and conclusive if the material properties could also cover a large range.

(2) The material property of the vocal fold was assumed to be linear based on the current understanding that the vocal fold vibration could be considered as small deformation. Though this assumption is also commonly used in other studies, it should be noted that in some conditions, the vibration amplitude is large that it might be more proper to use nonlinear material or include the geometric nonlinearity into the model. Moreover, as noticed in Chapter 3, with a linear material, the dependence of the fundamental frequency with subglottal pressure could not be observed.

In the future, this study could be extended in several aspects:

(1) The current partial differential equation based fluid solver could be applied in the study of unvoiced sound, where the vortex structures and turbulence in the glottal tract play an important role. The detailed flow field solution will help to understand the mechanism in this process.

(2) The current model could be used to improve the reduced order models. For example, the potential flow based flow solver has the potential in real time simulation but does not consider viscous loss and might fail in complex airway shapes such as the shapes with multi-channels. Partial differential equation based flow solver has the ability to provide accurate result in such condition. With proper combination with machine learning technologies, correction coefficients could be obtained for the reduced order flow models to provide more accurate solution of glottal flow rate and vocal fold dynamics.

BIBLIOGRAPHY

Adrian, R. J., Christensen, K. T. and Liu, Z. C. (2000) 'Analysis and interpretation of instantaneous turbulent velocity fields', *Experiments in Fluids*, 29(3), pp. 275–290. doi: 10.1007/s003489900087.

Agarwal, M. (2004) *The false vocal folds and their effect on translaryngeal airflow resistance*. Doctoral dissertation. Bowling Green State University.

Alipour-Haghighi, F., Perlman, A. L. and Titze, I. R. (1991) 'Tetanic response of the cricothyroid muscle', *Annals of Otology, Rhinology & Laryngology*, 100(8), pp. 626–631. doi: 10.1177/000348949110000805.

Alipour-Haghighi, F. and Titze, I. R. (1985) 'Simulation of particle trajectories of vocal fold tissue during phonation', *Vocal fold physiology: Biomechanics, acoustics and phonatory control*, pp. 183–190.

Alipour, F. *et al.* (2011) 'Mathematical Models and Numerical Schemes for the Simulation of Human Phonation', *Current Bioinformatics*, 43(0), pp. 323–343. doi: 10.2174/157489311796904655.

Alipour, F., Berry, D. A. and Titze, I. R. (2000) 'A finite-element model of vocal-fold vibration', *The Journal of the Acoustical Society of America*, 108(6), pp. 3003–3012. doi: 10.1121/1.1324678.

Alipour, F. and Jaiswal, S. (2008) 'Phonatory characteristics of excised pig, sheep, and cow larynges', *The Journal of the Acoustical Society of America*, 123(6), pp. 4572–4581. doi: 10.1121/1.2908289.

Alipour, F., Jaiswal, S. and Finnegan, E. (2007) 'Aerodynamic and acoustic effects of false vocal folds and epiglottis in excised larynx models', *Annals of Otology, Rhinology and Laryngology*, 116(2), pp. 135–144. doi: 10.1177/000348940711600210.

Alipour, F. and Scherer, R. C. (2000) 'Dynamic glottal pressures in an excised hemilarynx model', *Journal of Voice*, 14(4), pp. 443–454.

Alipour, F. and Scherer, R. C. (2004) 'Flow separation in a computational oscillating vocal fold model', *The Journal of the Acoustical Society of America*, 116(3), pp. 1710–1719. doi: 10.1121/1.1779274.

Alipour, F. and Scherer, R. C. (2012) 'Ventricular pressures in phonating excised larynges', *The Journal of the Acoustical Society of America*, 132(2), pp. 1017–1026. doi: 10.1121/1.4730880.

Bailly, L. *et al.* (2008) 'Influence of a constriction in the near field of the vocal folds: Physical modeling and experimental validation', *The Journal of the Acoustical Society of America*, 124(5), pp. 3296–3308. doi: 10.1121/1.2977740.

Bailly, L., Henrich, N. and Pelorson, X. (2010) 'Vocal fold and ventricular fold vibration in period-doubling phonation: Physiological description and aerodynamic modeling', *The Journal of the Acoustical Society of America*, 127(5), pp. 3212–3222. doi: 10.1121/1.3365220.

Baken, R. J. and Orlinkoff, R. F. (2000) *Clinical measurement of speech and voice*. 2nd edn, *Cengage Learning*. 2nd edn. doi: 10.1288/00005537-198808000-00028.

Bakhshaee, H. *et al.* (2013) 'Three-dimensional reconstruction of human vocal folds and standard laryngeal cartilages using computed tomography scan data', *Journal of Voice*, 27(6), pp. 769–777. doi: 10.1016/j.jvoice.2013.06.003.

Van den Berg, J. (1958) 'Myoelastic-aerodynamic theory of voice production.', *Journal of speech and hearing research*, 1(3), pp. 227–244. doi: 10.1044/jshr.0103.227.

Berke, G. *et al.* (2013) ' Neuromuscular induced phonation in a human ex vivo perfused larynx preparation ', *The Journal of the Acoustical Society of America*, 133(2), pp. EL114–EL117. doi: 10.1121/1.4776776.

Berry, D. A. *et al.* (1994) 'Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions.', *The Journal of the Acoustical Society of America*, 95(6), pp. 3595–3604. doi: 10.1121/1.409875.

Berry, D. A. *et al.* (1996) 'Bifurcations in excised larynx experiments', *Journal of Voice*, 10(2), pp. 129–138. doi: 10.1016/S0892-1997(96)80039-7.

Berry, D. A. (2001) 'Mechanism of modal and non-modal phonation', *Journal of Phonetics*, 29, pp. 431–450. doi: 10.1006/jpho.2001.0148.

Berry, D. A., Chhetri, D. K. and Neubauer, J. (2012) 'Observations of phonation threshold pressure and fundamental frequency in the in vivo canine larynx using graded neuromuscular stimulation', *The Journal of the Acoustical Society of America*, 131(4), pp. 3347–3347. doi: 10.1121/1.4708538.

Berry, D. A., Montequin, D. W. and Tayama, N. (2001) 'High-speed digital imaging of the medial surface of the vocal folds', *The Journal of the Acoustical Society of America*, 110(5), pp. 2539–2547. doi: 10.1121/1.1408947.

Bhattacharya, P., Kelleher, J. E. and Siegmund, T. (2015) 'Role of gradients in vocal fold elastic modulus on phonation', *Journal of Biomechanics*, 48(12), pp. 3356–3363. doi: 10.1016/j.jbiomech.2015.06.015.

Birk, V. *et al.* (2016) 'Acoustic impact of ventricular folds on phonation studied in ex vivo human larynx models', *Acta Acustica united with Acustica*, 102(2), pp. 244–256. doi: 10.3813/AAA.918941.

Bowling, D. L. *et al.* (2017) 'Body size and vocalization in primates and carnivores', *Scientific Reports*, 7. doi: 10.1038/srep41070.

Chan, R. W. and Titze, I. R. (1999) 'Viscoelastic shear properties of human vocal fold mucosa: measurement methodology and empirical results.', *Journal of the Acoustical Society of America*, 106(October 1999), pp. 2008–21. doi: 10.1121/1.427947.

Chan, R. W., Titze, I. R. and Titze, M. R. (1997) 'Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa', *The Journal of the Acoustical Society of America*, 101(6), pp. 3722–3727. doi: 10.1121/1.411870.

Chang, S. *et al.* (2016) 'Subject-Specific Computational Modeling of Evoked Rabbit Phonation', *Journal of Biomechanical Engineering*, 138(1), p. 011005. doi: 10.1115/1.4032057.

Chen, T. *et al.* (2012) 'A new method of reconstructing the human laryngeal architecture using micro-MRI', *Journal of Voice*, 26(5), pp. 555–562. doi: 10.1016/j.jvoice.2011.03.012.

Chhetri, D. K. and Park, S. J. (2016) 'Interactions of subglottal pressure and neuromuscular activation on fundamental frequency and intensity', *Laryngoscope*, 126(5), pp. 1123–1130. doi: 10.1002/lary.25550.

Chhetri, D. K. and Rafizadeh, S. (2014) 'Young's modulus of canine vocal fold cover layers', *Journal of Voice*, 28(4), pp. 406–410. doi: 10.1016/j.jvoice.2013.12.003.

Chhetri, D. K., Zhang, Z. and Neubauer, J. (2011) 'Measurement of Young's modulus of vocal folds by indentation', *Journal of Voice*, 25(1), pp. 1–7. doi: 10.1016/j.jvoice.2009.09.005.

Cisonni, J. *et al.* (2010) 'Experimental validation of quasi-one-dimensional and two-dimensional steady glottal flow models', *Medical and Biological Engineering and Computing*, 48(9), pp. 903–910. doi: 10.1007/s11517-010-0645-7.

Cook, D. D. A., Nauman, E. and Mongeau, L. (2008) 'Reducing the number of vocal fold mechanical tissue properties: evaluation of the incompressibility and planar displacement assumptions.', *The Journal of the Acoustical Society of America*, 124(6), pp. 3888–3896. doi: 10.1121/1.2996300.

Cook, D. D. and Mongeau, L. (2007) 'Sensitivity of a continuum vocal fold model to geometric parameters, constraints, and boundary conditions', *Journal of the Acoustical Society of America*, 121(4), pp. 2247–2253. doi: 10.1121/1.2536709.

Cook, D. D., Nauman, E. and Mongeau, L. (2009) 'Ranking vocal fold model parameters by their influence on modal frequencies', *The Journal of the Acoustical Society of America*, 126(4), p. 2002. doi: 10.1121/1.3183592.

Daily, D. J. and Thomson, S. L. (2013) 'Acoustically-coupled flow-induced vibration of a computational vocal fold model', *Computers and Structures*, 116, pp. 50–58. doi: 10.1016/j.compstruc.2012.10.022.

Decker, G. Z. and Thomson, S. L. (2007) 'Computational Simulations of Vocal Fold Vibration: Bernoulli Versus Navier-Stokes', *Journal of Voice*, 21(3), pp. 273–284. doi: 10.1016/j.jvoice.2005.12.002.

Dion, G. R. *et al.* (2016) 'Functional assessment of the ex vivo vocal folds through biomechanical testing: A review', *Materials science & engineering. C, Materials for biological applications*, 64, pp. 444–453. doi: 10.1016/j.msec.2016.04.018.

Dion, G. R. *et al.* (2019) 'Automated Indentation Mapping of Vocal Fold Structure and Cover Properties Across Species', *Laryngoscope*, 129(1), pp. E26–E31. doi: 10.1002/lary.27341.

Döllinger, M. *et al.* (2002) 'Vibration parameter extraction from endoscopic image series of the vocal folds', *IEEE Transactions on Biomedical Engineering*, 49(8), pp. 773–781. doi: 10.1109/TBME.2002.800755.

Döllinger, M., Berry, D. A. and Berke, G. S. (2005) ' Medial surface dynamics of an in vivo canine vocal fold during phonation ', *The Journal of the Acoustical Society of America*, 117(5), pp. 3174–3183. doi: 10.1121/1.1871772.

Dong, H., Mittal, R. and Najjar, F. M. (2006) 'Wake topology and hydrodynamic performance of low-aspect-ratio flapping foils', *Journal of Fluid Mechanics*, 566, pp. 309–343. doi: 10.1017/S002211200600190X.

Duncan, C., Zhai, G. and Scherer, R. C. (2006) 'Modeling coupled aerodynamics and vocal fold dynamics using immersed boundary methods', *The Journal of the Acoustical Society of America*, 120(5), pp. 2859–2871. doi: 10.1121/1.2354069.

Elemans, C. P. H. (2004) *How do birds sing? Sound analysis-mechanical modelling-muscular control*. Wageningen University.

Elemans, C. P. H. *et al.* (2015) 'Universal mechanisms of sound production and control in birds and mammals', *Nature Communications*, 6, p. 8978. doi: 10.1038/ncomms9978.

Elemans, C. P. H., Zaccarelli, R. and Herzel, H. (2008) 'Biomechanics and control of vocalization in a non-songbird', *Journal of the Royal Society Interface*, 5(24), pp. 691–703. doi: 10.1098/rsif.2007.1237.

Erath, B. D. *et al.* (2013) 'A review of lumped-element models of voiced speech', *Speech Communication*, 55(5), pp. 667–690. doi: 10.1016/j.specom.2013.02.002.

Erath, B. D. *et al.* (2019) 'An acoustic source model for asymmetric intraglottal flow with application to reduced-order models of the vocal folds', *PLoS ONE*, 14(7). doi: 10.1371/journal.pone.0219914.

Erath, B. D. and Plesniak, M. W. (2006a) 'An investigation of bimodal jet trajectory in flow through scaled models of the human vocal tract', *Experiments in Fluids*, 40(5), pp. 683–696. doi: 10.1007/s00348-006-0106-0.

Erath, B. D. and Plesniak, M. W. (2006b) 'An investigation of jet trajectory in flow through scaled vocal fold models with asymmetric glottal passages', *Experiments in Fluids*, 41(5), pp. 735–748. doi: 10.1007/s00348-006-0196-8.

Erath, B. D. and Plesniak, M. W. (2006c) 'The occurrence of the Coanda effect in pulsatile flow through static models of the human vocal folds', *The Journal of the Acoustical Society of America*, 120(2), pp. 1000–1011. doi: 10.1121/1.2213522.

Erath, B. D., Zañartu, M. and Peterson, S. D. (2017) 'Modeling viscous dissipation during vocal fold contact: the influence of tissue viscosity and thickness with implications for hydration', *Biomechanics and Modeling in Mechanobiology*, 16(3), pp. 947–960. doi: 10.1007/s10237-016-0863-5.

Fant G. (1960) 'Acoustic Theory of speech production', *The Hague*, pp. 1–24. doi: 10.2307/304731.

Farahani, M. H. *et al.* (2013) 'A numerical and experimental investigation of the effect of false vocal fold geometry on glottal flow.', *Journal of biomechanical engineering*, 135(12), p. 121006. doi: 10.1115/1.4025324.

Farahani, M. H. and Zhang, Z. (2014) 'A computational study of the effect of intraglottal vortex-induced negative pressure on vocal fold vibration', *The Journal of the Acoustical Society of America*, 136(5), pp. EL369–EL375. doi: 10.1121/1.4898743.

Farbos de Luzan, C. *et al.* (2015) 'Computational study of false vocal folds effects on unsteady airflows through static models of the human larynx', *Journal of Biomechanics*, 48(7), pp. 1248–1257. doi: 10.1016/j.jbiomech.2015.03.010.

Farbos de Luzan, C. *et al.* (2020) 'Volume velocity in a canine larynx model using time-resolved tomographic particle image velocimetry', *Experiments in Fluids*, 61(2). doi: 10.1007/s00348-020-2896-x.

Ffowcs Williams, J. and Hawkings, D. (1969) 'Sound Generation By Turbulence and Surfaces in Arbitrary Motion', *Roy Soc London-Philosophical Trans Ser A*, 264(1151), pp. 321–342. doi: 10.1098/rsta.1969.0031.

Flanagan, J. L. (1972) *Speech analysis : synthesis and perception*. New York: Springer.

Flanagan, J. L. and Landgraf, L. L. (1968) 'Self-oscillating source for vocal tract synthesizers', *IEEE Transactions on Audio and Electroacoustics*, 16(1), pp. 57–64.

Fulcher, L. P. and Scherer, R. C. (2019) 'Recent measurements with a synthetic two-layer model of the vocal folds and extension of Titze's surface wave model to a body-cover model', *The Journal of the Acoustical Society of America*, 146(6), pp. EL502–EL508. doi: 10.1121/1.5133664.

Fung, Y. C. (1993) *Biomechanics: Mechanical Properties of Living Tissues*. 2nd edn. New York: Springer-Verlag, New York. doi: 10.1115/1.3138285.

Garcia, M. and Herbst, C. T. (2018) 'Excised larynx experimentation: History, current developments, and prospects for bioacoustic research', *Anthropological Science*, 126(1), pp. 9–17. doi: 10.1537/ase.171216.

Geng, B., Xue, Q. and Zheng, X. (2016) 'The effect of vocal fold vertical stiffness variation on voice production', *The Journal of the Acoustical Society of America*, 140(4), pp. 2856–2866. doi: 10.1121/1.4964508.

Geng, B., Xue, Q. and Zheng, X. (2017) 'A finite element study on the cause of vocal fold vertical stiffness variation', *The Journal of the Acoustical Society of America*, 141(4), pp. EL351–EL356. doi: 10.1121/1.4978363.

Gómez, P. *et al.* (2017) 'Degrees of freedom in a vocal fold inverse problem', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10208 LNCS, pp. 475–484. doi: 10.1007/978-3-319-56148-6_42.

Gunter, H. E. (2004) 'Modeling mechanical stresses as a factor in the etiology of benign vocal fold lesions', *Journal of Biomechanics*, 37(7), pp. 1119–1124. doi: 10.1016/s0021-9290(03)00420-2.

Hatzikirou, H., Fitch, W. T. and Herzel, H. (2006) 'Voice instabilities due to source-tract interactions', *Acta Acustica united with Acustica*, 92(3), pp. 468–475.

Hirano, M. (1974) 'Morphological structure of the vocal cord as a vibrator and its variations', *Folia Phoniatrica et Logopaedica*, 26(2), pp. 89–94. doi: 10.1159/000263771.

Hirano, M. and Kakita, Y. (1985) 'Cover-body theory of vocal fold vibration', in *Speech Science*. San Diego, Ca: College-Hill Press, pp. 1–34.

Hirano, M., Kurita, S. and Nakashima, T. (1981) 'The Structure of the Vocal Folds', *Vocal Fold Physiology*, pp. 33–43.

Hirano, M., Kurita, S. and Sakaguchi, S. (1989) 'Ageing of the Vibratory Tissue of Human Vocal Folds', *Acta oto-laryngologica*, 107(5–6), pp. 428–433.

Holzapfel, G. G. A., Gasser, T. C. T. and Ogden, R. W. R. (2000) 'A new constitutive framework for arterial wall mechanics and a comparative study of material models', *J elasticity*, 61(8), pp. 1–48. doi: 10.1023/a:1010835316564.

Hunter, E. J., Titze, I. R. and Alipour, F. (2004) 'A three-dimensional model of vocal fold abduction/adduction', *The Journal of the Acoustical Society of America*, 115(4), pp. 1747–1759. doi: 10.1121/1.1652033.

Ishizaka, K. and Flanagan, J. L. (1972a) 'Synthesis of Voiced Sounds From a Two-Mass Model of the Vocal Cords', *Bell System Technical Journal*, 51(6), pp. 1233–1268. doi: 10.1002/j.1538-7305.1972.tb02651.x.

Ishizaka, K. and Flanagan, J. L. (1972b) 'Synthesis of Voiced Sounds From a Two-Mass Model of the Vocal Cords', *Bell System Technical Journal*, 51(6), pp. 1233–1268. doi: 10.1002/j.1538-7305.1972.tb02651.x.

Jiang, J. *et al.* (2004) 'Vocal efficiency measurements in subjects with vocal polyps and nodules: A preliminary report', *Annals of Otology, Rhinology and Laryngology*, 113(4), pp. 277–282.

Jiang, J. J., Diaz, C. E. and Hanson, D. G. (1998) 'Finite element modeling of vocal fold vibration in normal phonation and hyperfunctional dysphonia: Implications for the pathogenesis of vocal nodules', *Annals of Otology, Rhinology and Laryngology*, 107(7), pp. 603–610. doi: 10.1177/000348949810700711.

Jiang, W., Rasmussen, J. H., Xue, Q., Ding, M., Zheng, X., & Elemans, C. P. (2020) 'High-fidelity continuum modeling predicts avian voiced sound production', *Proceedings of the National Academy of Sciences of the United States of America*, 117(9), pp. 4718–4723. doi: 10.1073/pnas.1922147117.

Jiang, W., Xue, Q. and Zheng, X. (2018) 'Effect of longitudinal variation of vocal fold inner layer thickness on fluid-structure interaction during voice production', *Journal of Biomechanical Engineering*, 140(12). doi: 10.1115/1.4041045.

Jiang, W., Zheng, X. and Xue, Q. (2017) 'Computational modeling of fluid-structure-acoustics interaction during voice production', *Frontiers in Bioengineering and Biotechnology*, 5(FEB). doi: 10.3389/fbioe.2017.00007.

Jiang, W., Zheng, X. and Xue, Q. (2019) 'Influence of vocal fold cover layer thickness on its vibratory dynamics during voice production', *The Journal of the Acoustical Society of America*, 146(1), pp. 369–380. doi: 10.1121/1.5116567.

Jiao, Y. *et al.* (2018) 'Establishment and Analysis of False Vocal Folds Hypertrophy Model in Excised Canine Larynges', *Journal of Voice*, 32(2), pp. 143–148. doi: 10.1016/j.jvoice.2017.04.022.

van Kan, J. (1986) 'A Second-Order Accurate Pressure-Correction Scheme for Viscous Incompressible Flow', *SIAM Journal on Scientific and Statistical Computing*, 7(3), pp. 870–891. doi: 10.1137/0907059.

Kelleher, J. E. *et al.* (2010) 'Spatially varying properties of the vocal ligament contribute to its eigenfrequency response', *Journal of the Mechanical Behavior of Biomedical Materials*, 3(8), pp. 600–609. doi: 10.1016/j.jmbbm.2010.07.009.

Khosla, S. *et al.* (2008) 'Using particle imaging velocimetry to measure anterior-posterior velocity gradients in the excised canine larynx model', *Annals of Otology, Rhinology and Laryngology*, 117(2), pp. 134–144. doi: 10.1177/000348940811700212.

Kimura, M., Mau, T. and Chan, R. W. (2011) 'Rheometric properties of canine vocal fold tissues: Variation with anatomic location', *Auris Nasus Larynx*, 38(3), pp. 367–372. doi: 10.1016/j.anl.2010.09.006.

Kirmse, C. and Brücker, C. (2014) 'On the jet formation through a leaky glottis', *Journal of Fluids and Structures*, 50, pp. 137–152. doi: 10.1016/j.jfluidstructs.2014.06.022.

Kniesburges, S. *et al.* (2017) 'Effect of the ventricular folds in a synthetic larynx model', *Journal of Biomechanics*, 55, pp. 128–133. doi: 10.1016/j.jbiomech.2017.02.021.

Krebs, F. *et al.* (2012) 'A three-dimensional study of the glottal jet', *Experiments in Fluids*, 52(5), pp. 1133–1147. doi: 10.1007/s00348-011-1247-3.

Kucinschi, B. R. *et al.* (2006) 'Flow visualization and acoustic consequences of the air moving through a static model of the human larynx', *Journal of Biomechanical Engineering*, 128(3), pp. 380–390. doi: 10.1115/1.2187042.

LaMar, M. D., Qi, Y. and Xin, J. (2003) 'Modeling vocal fold motion with a hydrodynamic semicontinuum model', *The Journal of the Acoustical Society of America*, 114(1), pp. 455–464. doi: 10.1121/1.1577547.

Larsson, M. and Müller, B. (2009) 'Numerical simulation of confined pulsating jets in human phonation', *Computers and Fluids*, 38(7), pp. 1375–1383. doi: 10.1016/j.compfluid.2008.01.033.

Li, S. *et al.* (2006) 'The effect of glottal angle on intraglottal pressure', *The Journal of the Acoustical Society of America*, 119(1), pp. 539–548. doi: 10.1121/1.2133491.

Li, Z. *et al.* (2020) 'A Reduced-Order Flow Model for Fluid–Structure Interaction Simulation of Vocal Fold Vibration', *Journal of Biomechanical Engineering*, 142(2). doi: 10.1115/1.4044033.

Lighthill, M. J. (1982) 'On Sound Generated Aerodynamically. I. General Theory', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 205, pp. 581–598. doi: 10.1098/rspb.1979.0086.

Link, G. *et al.* (2009) 'A 2D finite-element scheme for fluid-solid-acoustic interactions and its application to human phonation', *Computer Methods in Applied Mechanics and Engineering*, 198(41–44), pp. 3321–3334. doi: 10.1016/j.cma.2009.06.009.

Liu, G. *et al.* (2019) 'An image-guided computational approach to inversely determine in vivo material properties and model flow-structure interactions of fish fins', *Journal of Computational Physics*, 392, pp. 578–593. doi: 10.1016/j.jcp.2019.04.062.

Lodermeyer, A. *et al.* (2018) 'Aeroacoustic analysis of the human phonation process based on a hybrid acoustic PIV approach', *Experiments in Fluids*, 59(1). doi: 10.1007/s00348-017-2469-9.

Ludlow, C. L. (2005) 'Central nervous system control of the laryngeal muscles in humans', *Respiratory Physiology and Neurobiology*, 147(2-3), pp. 205–222. doi: 10.1016/j.resp.2005.04.015.

Luo, H. *et al.* (2008) 'An immersed-boundary method for flow-structure interaction in biological systems with application to phonation', *Journal of Computational Physics*, 227(22), pp. 9303–9332. doi: 10.1016/j.jcp.2008.05.001.

Luo, H., Mittal, R. and Bielamowicz, S. A. (2009) 'Analysis of flow-structure interaction in the larynx during phonation using an immersed-boundary method.', *The Journal of the Acoustical Society of America*, 126(2), pp. 816–824. doi: 10.1121/1.3158942.

Manriquez, R. *et al.* (2019) 'Neurophysiological Muscle Activation Scheme for Controlling Vocal Fold Models', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(5), pp. 1043–1052. doi: 10.1109/TNSRE.2019.2906030.

Mattheus, W. and Brücker, C. (2011) 'Asymmetric glottal jet deflection: differences of two- and three-dimensional models.', *The Journal of the Acoustical Society of America*, 130(6), pp. EL373-9. doi: 10.1121/1.3655893.

Maxfield, L., Palaparthi, A. and Titze, I. (2016) 'New Evidence That Nonlinear Source-Filter Coupling Affects Harmonic Intensity and fo Stability During Instances of Harmonics Crossing Formants', *Journal of Voice*. doi: 10.1016/j.jvoice.2016.04.010.

McGowan, R. S. and Howe, M. S. (2010) 'ce with specified vocal fold motion', *The Journal of the Acoustical Society of America*, 127(3), pp. 1519–1527. doi: 10.1121/1.3299200.

McPhail, M. J., Campo, E. T. and Krane, M. H. (2019) 'Aeroacoustic source characterization in a physical model of phonation', *The Journal of the Acoustical Society of America*, 146(2), pp. 1230–1238. doi: 10.1121/1.5122787.

Mendelsohn, A. H. and Zhang, Z. (2011) 'Phonation threshold pressure and onset frequency in a two-layer physical model of the vocal folds', *The Journal of the Acoustical Society of America*, 130(5), pp. 2961–2968. doi: 10.1121/1.3644913.

Migimatsu, K. and Tokuda, I. T. (2019) 'Experimental study on nonlinear source–filter interaction using synthetic vocal fold models', *The Journal of the Acoustical Society of America*, 146(2), pp. 983–997. doi: 10.1121/1.5120618.

Mihaescu, M. *et al.* (2010) 'Unsteady laryngeal airflow simulations of the intra-glottal vortical structures', *The Journal of the Acoustical Society of America*, 127(1), pp. 435–444. doi: 10.1121/1.3271276.

Min, Y. B., Titze, I. R. and Alipour-Haghighi, F. (1995) 'Stress-strain response of the human vocal ligament', *Ann Otol Rhinol Laryngol*, 104(7), pp. 563–569. doi: 10.1177/000348949510400711.

Miri, A. K. (2014) 'Mechanical characterization of vocal fold tissue: A review study', *Journal of Voice*, 28(6), pp. 657–667. doi: 10.1016/j.jvoice.2014.03.001.

Mittal, R. *et al.* (2008) 'A versatile sharp interface immersed boundary method for incompressible flows with complex boundaries', *Journal of Computational Physics*, 227(10), pp. 4825–4852. doi: 10.1016/j.jcp.2008.01.028.

Moisik, S. R. and Esling, J. H. (2014) 'Modeling the biomechanical influence of epilaryngeal stricture on the vocal folds: A low-dimensional model of vocal-ventricular fold coupling', *Journal of Speech, Language, and Hearing Research*, 57(2), pp. 687–704. doi: 10.1044/2014_JSLHR-S-12-0279.

Moon, Y. J. *et al.* (2010) 'A hybrid prediction method for low-subsonic turbulent flow noise', *Computers and Fluids*, 39(7), pp. 1125–1135. doi: 10.1016/j.compfluid.2010.02.005.

Murray, P. R. and Thomson, S. L. (2011) 'Synthetic, Multi-Layer, Self-Oscillating Vocal Fold Model Fabrication', *Journal of Visualized Experiments*, (58). doi: 10.3791/3498.

Oren, L. *et al.* (2014) 'Characterization of the vocal fold vertical stiffness in a canine model', *Journal of Voice*, 28(3), pp. 297–304. doi: 10.1016/j.jvoice.2013.11.001.

Oren, L. *et al.* (2015) 'Intraglottal velocity and pressure measurements in a hemilarynx model', *The Journal of the Acoustical Society of America*, 137(2), pp. 935–943. doi: 10.1121/1.4906833.

Oren, L., Khosla, S. and Gutmark, E. (2014a) 'Intraglottal geometry and velocity measurements in canine larynges', *The Journal of the Acoustical Society of America*, 135(1), pp. 380–388. doi: 10.1121/1.4837222.

Oren, L., Khosla, S. and Gutmark, E. (2014b) 'Intraglottal pressure distribution computed from empirical velocity data in canine larynx', *Journal of Biomechanics*, 47(6), pp. 1287–1293. doi: 10.1016/j.jbiomech.2014.02.023.

Oren, L., Khosla, S. and Gutmark, E. (2019) 'Medial Surface Dynamics as a Function of Subglottal Pressure in a Canine Larynx Model', *Journal of Voice*. doi: 10.1016/j.jvoice.2019.07.015.

Pelorson, X. *et al.* (1994) 'Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. application to a modified two-mass model', *Journal of the Acoustical Society of America*, 96(6), pp. 3416–3431. doi: 10.1121/1.411449.

Pelorson, X. *et al.* (1995) 'Description of the flow through in-vitro models of the glottis during phonation', *Acta acustica (Les Ulis)*, 3(2), pp. 191–202.

Perrine, B. L. *et al.* (2020) 'Phonation threshold pressure using a 3-mass model of phonation with empirical pressure values', *The Journal of the Acoustical Society of America*, 147(3), pp. 1727–1737. doi: 10.1121/10.0000854.

Peskin, C. S. (1972) *Flow Patterns Around Heart Valve: A Digital Computer Method for Solving the Equations of Motion*. Albert Einstein College of Medicine.

Pham, N., Xue, Q. and Zheng, X. (2018) 'Coupling between a fiber-reinforced model and a Hill-based contractile model for passive and active tissue properties of laryngeal muscles: A finite element study', *The Journal of the Acoustical Society of America*, 144(3), pp. EL248–EL253. doi: 10.1121/1.5055564.

Pickup, B. A. and Thomson, S. L. (2011) 'Identification of geometric parameters influencing the flow-induced vibration of a two-layer self-oscillating computational vocal fold model', *The Journal of the Acoustical Society of America*, 129(4), pp. 2121–2132. doi: 10.1121/1.3557046.

Pirnia, A. *et al.* (2018) 'Discrete and periodic vortex loading on a flexible plate; application to energy harvesting and voiced speech production', *Journal of Sound and Vibration*, 433, pp. 476–492. doi: 10.1016/j.jsv.2018.05.046.

Rasmussen, J. H., Herbst, C. T. and Elemans, C. P. H. (2018) 'Quantifying syringeal dynamics in vitro using electroglottography', *Journal of Experimental Biology*, 221(16). doi: 10.1242/jeb.172247.

Rios, G. *et al.* (2019) 'Computational fluid dynamics analysis of surgical approaches to bilateral vocal fold immobility', *The Laryngoscope*. doi: 10.1002/lary.27925.

Rosa, M. D. O. *et al.* (2003) 'A contribution to simulating a three-dimensional larynx model using the finite element method.', *The Journal of the Acoustical Society of America*, 114(5), pp. 2893–2905. doi: 10.1121/1.1619981.

Sadeghi, H. *et al.* (2018) 'Computational Models of Laryngeal Aerodynamics: Potentials and Numerical Costs', *Journal of Voice*. doi: 10.1016/j.jvoice.2018.01.001.

Sadeghi, H., Döllinger, M., *et al.* (2019) 'Aerodynamic impact of the ventricular folds in computational larynx models', *The Journal of the Acoustical Society of America*, 145(4), pp. 2376–2387. doi: 10.1121/1.5098775.

Sadeghi, H., Kniesburges, S., *et al.* (2019) 'Towards a clinically applicable computational larynx model', *Applied Sciences (Switzerland)*, 9(11). doi: 10.3390/app9112288.

Scherer, R. C. *et al.* (2001) 'Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees', *The Journal of the Acoustical Society of America*, 109(4), pp. 1616–1630. doi: 10.1121/1.1333420.

Scherer, R. C. *et al.* (2010) 'Intraglottal pressures in a three-dimensional model with a non-rectangular glottal shape', *The Journal of the Acoustical Society of America*, 128(2), pp. 828–838. doi: 10.1121/1.3455838.

Sciamarella, D. and Le Quéré, P. (2008) 'Solving for unsteady airflow in a glottal model with immersed moving boundaries', *European Journal of Mechanics, B/Fluids*, 27(1), pp. 42–53. doi: 10.1016/j.euromechflu.2007.06.004.

Selbie, W. S., Gewalt, S. L. and Ludlow, C. L. (2002) 'Developing an anatomical model of the human laryngeal cartilages from magnetic resonance imaging', *The Journal of the Acoustical Society of America*, 112(3), pp. 1077–1090. doi: 10.1121/1.1501586.

Seo, J. H. and Mittal, R. (2011) 'A high-order immersed boundary method for acoustic wave scattering and low-Mach number flow-induced sound in complex geometries', *Journal of Computational Physics*, 230(4), pp. 1000–1019. doi: 10.1016/j.jcp.2010.10.017.

Seo, J. H. and Moon, Y. J. (2006) 'Linearized perturbed compressible equations for low Mach number aeroacoustics', *Journal of Computational Physics*, 218(2), pp. 702–719. doi: 10.1016/j.jcp.2006.03.003.

Seo, J. H. and Moon, Y. J. (2007) 'Aerodynamic noise prediction for long-span bodies', *Journal of Sound and Vibration*, 306(3–5), pp. 564–579. doi: 10.1016/j.jsv.2007.05.042.

Šidlof, P. *et al.* (2011) 'Measurement of flow separation in a human vocal folds model', *Experiments in Fluids*, 51(1), pp. 123–136. doi: 10.1007/s00348-010-1031-9.

Šidlof, P., Zörner, S. and Hüppe, A. (2013) 'Numerical simulation of flow-induced sound in human voice production', *Procedia Engineering*, 61, pp. 333–340. doi: 10.1016/j.proeng.2013.08.024.

Šidlof, P., Zörner, S. and Hüppe, A. (2014) 'A hybrid approach to the computational aeroacoustics of human voice production', *Biomechanics and Modeling in Mechanobiology*, 14(3), pp. 473–488. doi: 10.1007/s10237-014-0617-1.

Smith, S. L. and Hunter, E. J. (2014) 'A viscoelastic laryngeal muscle model with active components', *The Journal of the Acoustical Society of America*, 135(4), pp. 2041–2051. doi: 10.1121/1.4866173.

Smith, S. L. and Thomson, S. L. (2012) 'Effect of inferior surface angle on the self-oscillation of a computational vocal fold model', *The Journal of the Acoustical Society of America*, 131(5), pp. 4062–4075. doi: 10.1121/1.3695403.

Sondhi, M. M. and Schroeter, J. (1987) 'A Hybrid Time-Frequency Domain Articulatory Speech Synthesizer', *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(7), pp. 955–967. doi: 10.1109/TASSP.1987.1165240.

Sparrow, E. M., Abraham, J. P. and Minkowycz, W. J. (2009) 'Flow separation in a diverging conical duct: Effect of Reynolds number and divergence angle', *International Journal of Heat and Mass Transfer*, 52(13–14), pp. 3079–3083. doi: 10.1016/j.ijheatmasstransfer.2009.02.010.

Steinecke, I. and Herzel, H. (1995) 'Bifurcations in an asymmetric vocal-fold model', *Journal of the Acoustical Society of America*, 97(3), pp. 1874–1884. doi: 10.1121/1.412061.

Stewart, K. C., Erath, B. D. and Plesniak, M. W. (2014) 'Investigating the three-dimensional flow separation induced by a model vocal fold polyp', *Journal of Visualized Experiments*, (84). doi: 10.3791/51080.

Story, B. H. (2005) 'A parametric model of the vocal tract area function for vowel and consonant simulation', *The Journal of the Acoustical Society of America*, 117(5), pp. 3231–3254. doi: 10.1121/1.1869752.

Story, B. H. and Titze, I. R. (1995) 'Voice simulation with a body-cover model of the vocal folds', *Journal of the Acoustical Society of America*, 97(2), pp. 1249–1260. doi: 10.1121/1.412234.

Story, B. H., Titze, I. R. and Hoffman, E. A. (1996) 'Vocal tract area functions from magnetic resonance imaging', *J Acoust Soc Am*, 100(1), pp. 537–554. doi: Doi 10.1121/1.415960.

Story, B. H. and Titze, I. R. (1998) 'Parametrization of vocal tract area functions by empirical orthogonal modes', *Journal of Phonetics*, 26(November), pp. 223–260. doi: http://dx.doi.org/10.1006/jpho.1998.0076.

Suh, J. and Frankel, S. H. (2007) 'Numerical simulation of turbulence transition and sound radiation for flow through a rigid glottal model.', *The Journal of the Acoustical Society of America*, 121(6), pp. 3728–3739. doi: 10.1121/1.2723646.

Suh, J. and Frankel, S. H. (2008) 'Comparing turbulence models for flow through a rigid glottal model', *The Journal of the Acoustical Society of America*, 123(3), pp. 1237–1240. doi: 10.1121/1.2836783.

Sundström, E. and Oren, L. (2019) 'Pharyngeal flow simulations during sibilant sound in a patient-specific model with velopharyngeal insufficiency', *The Journal of the Acoustical Society of America*, 145(5), pp. 3137–3145. doi: 10.1121/1.5108889.

Tang, J. and Stathopoulos, E. T. (1995) 'Vocal efficiency as a function of vocal intensity: A study of children, women, and men', *Journal of the Acoustical Society of America*, 97(3), pp. 1885–1892. doi: 10.1121/1.412062.

Tao, C. *et al.* (2007) 'Asymmetric airflow and vibration induced by the Coanda effect in a symmetric model of the vocal folds.', *The Journal of the Acoustical Society of America*, 122(4), pp. 2270–8. doi: 10.1121/1.2773960.

Tao, C. and Jiang, J. J. (2007) 'Mechanical stress during phonation in a self-oscillating finite-element vocal fold model', *Journal of Biomechanics*, 40(10), pp. 2191–2198. doi: 10.1016/j.jbiomech.2006.10.030.

Thomson, S. L., Mongeau, L. and Frankel, S. H. (2005) 'Aerodynamic transfer of energy to the vocal folds', *The Journal of the Acoustical Society of America*, 118(3), p. 1689. doi: 10.1121/1.2000787.

Titze, I. R. (1973) 'The Human Vocal Cords: A Mathematical Model', *Phonetica*, 28, pp. 129–170. doi: 10.1159/000259461.

Titze, I. R. (1988) 'The physics of small-amplitude oscillation of the vocal folds', *Journal of the Acoustical Society of America*, 83(4), pp. 1536–1552. doi: 10.1121/1.395910.

Titze, I. R. (1992a) 'Phonation threshold pressure: A missing link in glottal aerodynamics', *Journal of the Acoustical Society of America*, 91(5), pp. 2926–2935. doi: 10.1121/1.402928.

Titze, I. R. (1992b) 'Vocal efficiency', *Journal of Voice*, 6(2), pp. 135–138. doi: 10.1016/S0892-1997(05)80127-4.

Titze, I. R. (2000) *Principles of Voice Production*. Iowa, City, IA: National Center for Voice and Speech, Iowa City, IA.

Titze, I. R. (2006a) *The myoelastic aerodynamic theory of phonation*. Iowa City IA: National Center for Voice and Speech.

Titze, I. R. (2006b) 'Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation', *J Speech Lang Hear Res*, 49(2), pp. 439–447. doi: 10.1044/1092-4388(2006/034).

Titze, I. R. (2008) 'Nonlinear source-filter coupling in phonation: Theory', *Journal of the Acoustical Society of America*, 123(5), pp. 2733–2749. doi: 10.1121/1.2832337.

Titze, I. R. (2011) 'Vocal Fold Mass Is Not A Useful Quantity for Describing F0 in Vocalization', *Journal of Speech Language and Hearing Research*, 54(2), pp. 520–522. doi: 10.1044/1092-4388(2010/09-0284).

Titze, I. R., Jiang, J. and Drucker, D. G. (1988) 'Preliminaries to the body-cover theory of pitch control', *Journal of Voice*, 1(4), pp. 314–319. doi: 10.1016/S0892-1997(88)80004-3.

Titze, I. R., Jiang, J. J. and Hsiao, T. yu (1993) 'Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration', *Annals of Otology, Rhinology & Laryngology*, 102(1), pp. 58–63. doi: 10.1177/000348949310200111.

Titze, I. R., Maxfield, L. and Palaparthi, A. (2016) 'An Oral Pressure Conversion Ratio as a Predictor of Vocal Efficiency', *Journal of Voice*, 30(4), pp. 398–406. doi: 10.1016/j.jvoice.2015.06.002.

Titze, I. R., Schmidt, S. S. and Titze, M. R. (1995) 'Phonation threshold pressure in a physical model of the vocal fold mucosa', *Journal of the Acoustical Society of America*, 97(5), pp. 3080–3084. doi: 10.1121/1.411870.

Titze, I. R. and Story, B. H. (2002) 'Rules for controlling low-dimensional vocal fold models with muscle activation', *The Journal of the Acoustical Society of America*, 112(3), pp. 1064–1076. doi: 10.1121/1.1496080.

Titze, I. R. and Talkin, D. T. (1979) 'A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation', *The Journal of the Acoustical Society of America*, 66(1), pp. 60–74. doi: 10.1121/1.382973.

Tokuda, I. T. *et al.* (2007) 'Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments', *The Journal of the Acoustical Society of America*, 122(1), pp. 519–531. doi: 10.1121/1.2741210.

Tokuda, I. T. *et al.* (2008) 'Bifurcations and chaos in register transitions of excised larynx experiments', *Chaos*, 18(1). doi: 10.1063/1.2825295.

Towns, J. *et al.* (2014) 'XSEDE: Accelerating scientific discovery', *Computing in Science and Engineering*, 16(5), pp. 62–74. doi: 10.1109/MCSE.2014.80.

Triep, M. and Brücker, C. (2010) 'Three-dimensional nature of the glottal jet', *The Journal of the Acoustical Society of America*, 127(3), pp. 1537–1547. doi: 10.1121/1.3299202.

Tse, J. R., Zhang, Z. and Long, J. L. (2015) 'Effects of vocal fold epithelium removal on vibration in an excised human larynx model', *The Journal of the Acoustical Society of America*, 138(1), pp. EL60–EL64. doi: 10.1121/1.4922765.

Unteregger, F. *et al.* (2018) 'Changes in Vocal Fold Morphology During Singing Over Two Octaves', *Journal of Voice*. doi: 10.1016/j.jvoice.2018.08.020.

Vahabzadeh-Hagh, A. M., Zhang, Z. and Chhetri, D. K. (2017a) 'Quantitative Evaluation of the In Vivo Vocal Fold Medial Surface Shape', *Journal of Voice*, 31(4), pp. 513.e15-513.e23. doi: 10.1016/j.jvoice.2016.12.004.

Vahabzadeh-Hagh, A. M., Zhang, Z. and Chhetri, D. K. (2017b) 'Three-dimensional posture changes of the vocal fold from paired intrinsic laryngeal muscles', *Laryngoscope*, 127(3), pp. 656–664. doi: 10.1002/lary.26145.

Wang, X. *et al.* (2019) 'A computational study of the effects of vocal fold stiffness parameters on voice production', *Journal of Voice*. doi: 10.1016/j.jvoice.2019.09.004.

Weiss, S. *et al.* (2016) 'Measurement and analysis of the material properties and oscillation characteristics of synthetic vocal folds', *Acta Acustica united with Acustica*, 102(2), pp. 214–229. doi: 10.3813/AAA.918938.

Wong, D. *et al.* (1991) 'Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases', *Journal of the Acoustical Society of America*, 89(1), pp. 383–394. doi: 10.1121/1.400472.

Wu, L. and Zhang, Z. (2016) 'A parametric vocal fold model based on magnetic resonance imaging', *The Journal of the Acoustical Society of America*, 140(2), pp. EL159–EL165. doi: 10.1121/1.4959599.

Wu, L. and Zhang, Z. (2019) 'Voice production in a MRI-based subject-specific vocal fold model with parametrically controlled medial surface shape', *The Journal of the Acoustical Society of America*, 146(6), pp. 4190–4198. doi: 10.1121/1.5134784.

Xuan, Y. and Zhang, Z. (2014) 'Influence of embedded fibers and an epithelium layer on the glottal closure pattern in a physical vocal fold model.', *Journal of speech, language, and hearing research : JSLHR*, 57(2), pp. 416–425. doi: 10.1044/2013.

Xue, Q. *et al.* (2010) 'A computational study of the effect of vocal-fold asymmetry on phonation.', *The Journal of the Acoustical Society of America*, 128(2), pp. 818–827. doi: 10.1121/1.3458839.

Xue, Q. (2011) *Computational Modeling and Analysis of Fluid-Structure Interaction During Phonation in Healthy and Diseased Larynges*. Doctoral dissertation. The Johns Hopkins University.

Xue, Q. *et al.* (2011) 'Sensitivity of vocal fold vibratory modes to their three-layer structure: Implications for computational modeling of phonation', *The Journal of the Acoustical Society of America*, 130(2), pp. 965–976. doi: 10.1121/1.3605529.

Xue, Q. *et al.* (2012) 'Computational modeling of phonatory dynamics in a tubular three-dimensional model of the human larynx', *The Journal of the Acoustical Society of America*, 132(3), pp. 1602–1613. doi: 10.1121/1.4740485.

Xue, Q. *et al.* (2014) 'Subject-specific computational modeling of human phonation', *The Journal of the Acoustical Society of America*, 135(3), pp. 1445–1456. doi: 10.1121/1.4864479.

Xue, Q. and Zheng, X. (2017) 'The Effect of False Vocal Folds on Laryngeal Flow Resistance in a Tubular Three-dimensional Computational Laryngeal Model', *Journal of Voice*, 31(3), pp. 275–281. doi: 10.1016/j.jvoice.2016.04.006.

Yang, J. *et al.* (2017). 'Fully-coupled aeroelastic simulation with fluid compressibility—For application to vocal fold vibration', *Computer methods in applied mechanics and engineering*, 315, pp. 584-606. doi: 10.1016/j.cma.2016.11.010.

Yin, J. and Zhang, Z. (2013) 'The influence of thyroarytenoid and cricothyroid muscle activation on vocal fold stiffness and eigenfrequencies', *The Journal of the Acoustical Society of America*, 133(5), pp. 2972–2983. doi: 10.1121/1.4799809.

Yin, J. and Zhang, Z. (2014) 'Interaction between the thyroarytenoid and lateral cricoarytenoid muscles in the control of vocal fold adduction and eigenfrequencies', *Journal of Biomechanical Engineering*, 136(11). doi: 10.1115/1.4028428.

Yin, J. and Zhang, Z. (2016) 'Laryngeal muscular control of vocal fold posturing: Numerical modeling and experimental validation', *The Journal of the Acoustical Society of America*, 140(3), pp. EL280–EL284. doi: 10.1121/1.4962375.

Yoshinaga, T., Nozaki, K. and Wada, S. (2019) 'Aeroacoustic analysis on individual characteristics in sibilant fricative production', *The Journal of the Acoustical Society of America*, 146(2), pp. 1239–1251. doi: 10.1121/1.5122793.

Zañartu, M. *et al.* (2011) 'Observation and analysis of in vivo vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study', *The Journal of the Acoustical Society of America*, 129(1), pp. 326–339. doi: 10.1121/1.3514536.

Zañartu, M., Mongeau, L. and Wodicka, G. R. (2007) 'Influence of acoustic loading on an effective single mass model of the vocal folds.', *The Journal of the Acoustical Society of America*, 121(2), pp. 1119–1129. doi: 10.1121/1.2409491.

Zhang, C. *et al.* (2002) 'Computational aeroacoustics of phonation, Part II: Effects of flow parameters and ventricular folds', *The Journal of the Acoustical Society of America*, 112(5), pp. 2147–2154. doi: 10.1121/1.1506694.

Zhang, K., Siegmund, T. and Chan, R. W. (2009) 'Modeling of the transient responses of the vocal fold lamina propria', *Journal of the Mechanical Behavior of Biomedical Materials*, 2(1), pp. 93–104. doi: 10.1016/j.jmbbm.2008.05.005.

Zhang, L. T., Krane, M. H. and Yu, F. (2019) 'Modeling of slightly-compressible isentropic flows and compressibility effects on fluid-structure interactions', *Computers and Fluids*, 182, pp. 108–117. doi: 10.1016/j.compfluid.2019.02.013.

Zhang, Y., Zheng, X. and Xue, Q. (2020) 'A deep neural network based glottal flow model for predicting fluid-structure interactions during voice production', *Applied Sciences (Switzerland)*, 10(2). doi: 10.3390/app10020705.

Zhang, Z. (2008) 'Influence of flow separation location on phonation onset.', *The Journal of the Acoustical Society of America*, 124(3), pp. 1689–94. doi: 10.1121/1.2957938.

Zhang, Z. (2009) 'Characteristics of phonation onset in a two-layer vocal fold model.', *The Journal of the Acoustical Society of America*, 125(2), pp. 1091–102. doi: 10.1121/1.3050285.

Zhang, Z. (2010) 'Dependence of phonation threshold pressure and frequency on vocal fold geometry and biomechanics', *The Journal of the Acoustical Society of America*, 127(4), pp. 2554–2562. doi: 10.1121/1.3308410.

Zhang, Z. (2014) 'The influence of material anisotropy on vibration at onset in a three-dimensional vocal fold model', *The Journal of the Acoustical Society of America*, 135(3), pp. 1480–1490. doi: 10.1121/1.4863266.

Zhang, Z. (2015) 'Regulation of glottal closure and airflow in a three-dimensional phonation model: Implications for vocal intensity control.', *Journal of the Acoustical Society of America*, 137(2), pp. 898–910. doi: 10.1121/1.4906272.

Zhang, Z. (2016a) 'Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model', *The Journal of the Acoustical Society of America*, 139(4), pp. 1493–1507. doi: 10.1121/1.4944754.

Zhang, Z. (2016b) 'Mechanics of human voice production and control', *The Journal of the Acoustical Society of America*, 140(4), pp. 2614–2635. doi: 10.1121/1.4964509.

Zhang, Z. (2017a) 'Effect of vocal fold stiffness on voice production in a three-dimensional body-cover phonation model', *The Journal of the Acoustical Society of America*, 142(4), pp. 2311–2321. doi: 10.1121/1.5008497.

Zhang, Z. (2017b) 'Toward real-time physically-based voice simulation: An Eigenmode-based approach', *Proceedings of Meetings on Acoustics*, 30(1). doi: 10.1121/2.0000572.

Zhang, Z. (2019) 'Vocal fold contact pressure in a three-dimensional body-cover phonation model', *The Journal of the Acoustical Society of America*, 146(1), pp. 256–265. doi: 10.1121/1.5116138.

Zhang, Z. and Luu, T. H. (2012) 'Asymmetric vibration in a two-layer vocal fold model with left-right stiffness asymmetry: Experiment and simulation', *The Journal of the Acoustical Society of America*, 132(3), pp. 1626–1635. doi: 10.1121/1.4739437.

Zhang, Z., Mongeau, L. and Frankel, S. H. (2002) 'Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes', *The Journal of the Acoustical Society of America*, 112(4), pp. 1652–1663. doi: 10.1121/1.1506159.

Zhang, Z. and Mongeau, L. G. (2006) 'Broadband sound generation by confined pulsating jets in a mechanical model of the human larynx', *The Journal of the Acoustical Society of America*, 119(6), pp. 3995–4005. doi: 10.1121/1.2195268.

Zhang, Z., Neubauer, J. and Berry, D. A. (2009) 'Influence of vocal fold stiffness and acoustic loading on flow-induced vibration of a single-layer vocal fold model', *Journal of Sound and Vibration*, 322(1–2), pp. 299–313. doi: 10.1016/j.jsv.2008.11.009.

Zhang, Z., Neubauer, J. and Berry, D. A. (2006a) 'Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds.', *The Journal of the Acoustical Society of America*, 120, pp. 2841–2849. doi: 10.1121/1.2354025.

Zhang, Z., Neubauer, J. and Berry, D. A. (2006b) 'The influence of subglottal acoustics on laboratory models of phonation', *Journal of the Acoustical Society of America*, 120(3), pp. 1558–1569. doi: 10.1121/1.2225682.

Zhang, Z., Neubauer, J. and Berry, D. A. (2007) 'Physical mechanisms of phonation onset: a linear stability analysis of an aeroelastic continuum model of phonation.', *The Journal of the Acoustical Society of America*, 122(October), pp. 2279–2295. doi: 10.1121/1.2773949.

Zhao, W. *et al.* (2002) 'Computational aeroacoustics of phonation, part I: Computational methods and sound generation mechanisms', *Journal of the Acoustical Society of America*, 112(5), pp. 2134–2146. doi: 10.1121/1.1506693.

Zhao, W., Frankel, S. H. and Mongeau, L. (2001) 'Numerical simulations of sound from confined pulsating axisymmetric jets', *AIAA Journal*, 39(10), pp. 1868–1874. doi: 10.2514/3.14944.

Zheng, X. *et al.* (2009) 'A computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation', *Annals of Biomedical Engineering*, 37(3), pp. 625–642. doi: 10.1007/s10439-008-9630-9.

Zheng, X. (2009) *Biomechanical Modeling of Glottal Aerodynamics and Vocal Fold Vibration During Phonation*. Doctoral dissertation. George Washington University, DC.

Zheng, X. *et al.* (2010) 'A Coupled Sharp-Interface Immersed Boundary-Finite-Element Method for Flow-Structure Interaction With Application to Human Phonation', *Journal of Biomechanical Engineering*, 132(11), p. 111003. doi: 10.1115/1.4002587.

Zheng, X. *et al.* (2011) 'Direct-numerical simulation of the glottal jet and vocal-fold dynamics in a three-dimensional laryngeal model', *The Journal of the Acoustical Society of America*, 130(1), pp. 404–415. doi: 10.1121/1.3592216.

Zheng, X., Mittal, R. and Bielamowicz, S. (2011) 'A computational study of asymmetric glottal jet deflection during phonation', *The Journal of the Acoustical Society of America*, 129(4), pp. 2133–2143. doi: 10.1121/1.3544490.

BIOGRAPHY OF THE AUTHOR

Weili Jiang was born in Nanjing, Jiangsu, China in 1990. She was also raised in Nanjing. She earned the Bachelor's degree in the Southeast University from 2008 to 2012 with the major of Thermal Energy and Power Engineering and the Master's degree from 2012 to 2015 with the major of Power Engineering and Engineering Thermophysics. Later, she joined the PhD program of University of Maine in the Mechanical department in May 2015. She is a candidate for the Doctor of Philosophy degree in Mechanical Engineering from the University of Maine in August 2020.